CHEMICAL REVIEWS

VOLUME 26

PUBLISHED BI-MONTHLY FOR
The American Chemical Society
BY
The Williams & Wilkins Company
Baltimore, U. S. A.
1940

CONTENTS

Number 1, February, 1940

Number 2, April, 1940

Number 3, June, 1940

SYMPOSIUM ON THE CHEMISTRY OF LIQUID AMMONIA SOLUTIONS[1]

INTRODUCTION TO THE SYMPOSIUM

WARREN C. JOHNSON

Department of Chemistry, University of Chicago, Chicago, Illinois

Received September 27, 1939

Researches concerning liquid ammonia solutions were initiated in this country about forty years ago at the University of Kansas. The pioneers in this field were Cady, Franklin, and Kraus, all of whom are well known in American chemistry for their contributions to the science. Much of the earlier work on the physical properties of ammonia solutions is due to Cady. Franklin not only contributed to our knowledge of the physical and chemical properties of these solutions but, in addition, developed a rational system for nitrogen compounds, which has enabled workers to classify and predict reactions on the basis of the more familiar reactions of oxygen compounds in an aqueous medium. Chemistry as a whole, and especially liquid ammonia chemistry, has recently suffered a great loss in the passing of Franklin. The absence of his scientific contributions, as well as his unusually fine personality, is greatly felt by American chemistry.

Kraus' interest in ammonia solutions has been sustained over this period of more than forty years. His earlier work was primarily concerned with the physical properties of ammonia solutions, with special reference to the ionization theory and the properties of metals. He later became more interested in the chemistry of ammonia solutions, since it was early recognized that ammonia is a very suitable solvent for studying the chemistry of many types of compounds, some of which are incapable of existence in an aqueous medium.

It was also recognized at a rather early date that liquid ammonia is an excellent solvent for organic compounds; as a matter of fact, it appears to resemble alcohol more than any other solvent in its ability to dissolve

[1] This Symposium was held under the joint auspices of the Division of Physical and Inorganic Chemistry and the Division of Organic Chemistry at the Ninety-sixth Meeting of the American Chemical Society, held at Milwaukee, Wisconsin, September, 1938.

organic substances. Thus, it is not surprising that in the past two decades a great deal of attention has been given to the chemistry of ammonia solutions of organic compounds.

Between the fields of organic chemistry and inorganic chemistry is another, that of the metallo-organic compounds. A considerable portion of our present knowledge of the chemistry of these compounds has been gained through studies in which liquid ammonia has served as the solvent medium.

Accordingly, it is more than appropriate, it is essential that organic and inorganic chemists get together to present results of their investigations of ammonia solutions. It was with this purpose in mind that this symposium was organized.

AMMONOLYSIS IN LIQUID AMMONIA[1]

W. CONARD FERNELIUS AND GLADE B. BOWMAN

Department of Chemistry, The Ohio State University, Columbus, Ohio

Received August 3, 1939

CONTENTS

[1] This paper was presented in abstract form at the Symposium on the Chemistry of Liquid Ammonia Solutions, held at Milwaukee, Wisconsin, September, 1938, under the auspices of the Division of Physical and Inorganic Chemistry and the Division of Organic Chemistry of the American Chemical Society.

I. Introduction

The term "ammonolysis" (75, 88) is used to designate those reactions of ammonia which are analogous to the familiar hydrolysis reactions of water. Thus an ammonolytic reaction is one of double decomposition in which ammonia is a reactant. Since substances having acid and basic characteristics are formed in such reactions, ammonolytic reactions constitute one type of protolytic equilibria. Numerous ammonolytic reactions are known, and certain of them are widely used in preparative chemistry. To one interested in reactions in liquid ammonia the possibility of ammonolysis is always present and must be taken into consideration at all times. In many cases ammonolytic reactions may be turned to good advantage; in others, ammonolysis acts to prevent one from realizing a desired preparation. Further, the study of certain ammonolytic reactions offers considerable promise in providing a better understanding of the physical-chemical characteristics of liquid ammonia.

One is naturally interested in knowing wherein ammonolysis differs from hydrolysis. Ammonolysis, like hydrolysis, is dependent upon the degree of auto-ionization of the solvent,

$$2NH_3 \rightleftharpoons NH_4^+ + NH_2^-$$

The character of many reactions in liquid ammonia as well as the low specific conductance of this solvent indicate that the ion-product of ammonia,

$$K = [NH_4^+]\,[NH_2^-]$$

is much less than the corresponding constant for water. However, the

quantitative measurement of this constant is difficult. Because of the difficulty of purifying ammonia, calculations based on the specific conductance and the independently determined conductivities of the ammonium and amide ions give a value for K that is much too large. The best value for this constant,

$$K = 1.9 \times 10^{-33} \ (-50°C.)$$

is that of Pleskov and Monoszon (223), who based their calculations on the E.M.F. of acid–base cells of the type,

$$\begin{array}{llll} Pt, H_2 \mid NH_4NO_3 & KNO_3 & KNH_2 & Pt, H_2 \\ \quad\ 0.1 \text{ normal} & \text{saturated} & & \end{array}$$

The stability of solutions of metals in liquid ammonia and the lack of ammonolysis of aluminum cyanide solutions (9) are undoubtedly due to the low self-ionization of liquid ammonia.

Many hydrolytic reactions, especially those of esters, are known to be subject to acid–base catalysis (39, 40, 49, 50, 67, 111, 127, 146). Several years ago, Franklin and coworkers (26, 47, 86, 264) performed a number of experiments which indicated that ammonolytic reactions were acid-catalyzed. More recently the ammonolysis of a number of esters has been studied in a quantitative manner (5, 56, 240–1, 243–4, 246, 251). At this point one would expect to find considerable difference between the behavior of water and ammonia because of the inherent greater basicity of the latter (7, 109, 110, 242). Owing to the more pronounced "leveling effect" of ammonia upon acidity, it is impossible to obtain as strongly acid solutions in ammonia as in water. In other words, the acid strength and, consequently, the catalytic activity of the ammonium ion is less than that of the hydronium ion.

Although ammonolytic reactions have been extensively studied, comparatively little work has been directed toward establishing definitely the influence of conditions upon the nature and extent of a reaction. The techniques commonly used for studying reactions in liquid ammonia may be divided into two major classes: (*1*) reactions at low temperatures (up to −33°C.) and at atmospheric pressure or less and (*2*) reactions under pressure: (*a*) in sealed glass tubes (temperatures up to about 80°C.) and (*b*) in autoclaves (temperatures up to and beyond the critical temperature, 132°C., of ammonia) (*cf.* reference 58 for bibliography). Since the results of any study on ammonolysis will vary, depending upon the particular technique used, the experimental conditions are given in this review wherever they are known.

The extent to which ammonolysis will proceed in any instance will depend upon (*1*) the ion-product of liquid ammonia, (*2*) the relative strengths

of the acid and base concerned, (*3*) the ratio of ammonia to substance ammonolyzed, and (*4*) the solubilities of the substance undergoing ammonolysis and of the ammonolytic products. Since the ion-product of ammonia is undoubtedly increased with rise in temperature (7), the degree of ammonolysis will also depend upon the temperature. Finally, ammonolysis may be brought to completion by removing one of the substances formed by the reaction. Thus, providing the solubility relationships are suitable, simple extraction of a soluble product with liquid ammonia will bring the reaction to completion, or ammonium ion may be destroyed by adding amide ion or a soluble metal

$$NH_4^+ + NH_2^- \rightarrow 2NH_3$$

$$NH_4^+ + e^- \rightarrow NH_3 + \tfrac{1}{2}H_2$$

and amide ion may be destroyed by adding ammonium ion.

Like water, many of the alcohols enter into solvolytic reactions. Similarly, the simple primary and secondary amines exhibit reactions of *aminolysis* (*cf.* 59). For any given substance, the extent of aminolysis is less than that of ammonolysis. Consequently, in many preparations where the use of liquid ammonia is unsuited because of extensive ammonolysis, the difficulties may be overcome (provided the solubility relationships are favorable) by using a simple amine such as ethylamine (156, 161, 165).

II. Ammonolysis of Inorganic Compounds

A. INTRODUCTION

Although there have been numerous studies on the reactions of ammonia with inorganic substances, particularly halides, much of this work has not been conducted in such a manner as to enable one to state definitely whether or not ammonolysis has taken place. This situation arises because many investigators have failed to distinguish between *ammonolysis* and *ammonation*. The literature of chemistry contains many references to "ammonates" prepared by treating a halide, either in the pure state or in solution in an inert solvent, with ammonia. The solid obtained in such cases is not of necessity an ammonate but may be a mixture of an amide, imide, nitride, or ammonobasic salt and ammonium halide. To illustrate, on treatment with ammonia germanic iodide gains in weight by an amount corresponding to 6 NH_3. On the basis of this fact *alone*, one may represent the resulting solid equally well as $GeI_4 \cdot 6NH_3$ or $Ge(NH)_2 + 4NH_4I$. In several cases, the extraction of such mixtures with liquid ammonia has demonstrated the existence of ammonolytic products (34, 137, 140, 236, 259, 270).

B. HYDRIDES

The alkali-metal hydrides react with liquid ammonia, yielding hydrogen and metal amides (229; *cf.* 3).

$$MH + NH_3 \rightarrow MNH_2 + H_2$$

With sodium hydride, the reaction proceeds slowly at −40°C. and rapidly at 20°C. The volatile hydrides of the fourth and fifth groups of the periodic table (carbon, silicon, germanium, phosphorus, arsenic) are unreactive toward ammonia, while those of the sixth and seventh groups (oxygen, sulfur, selenium, fluorine, chlorine, bromine, iodine) form ammonium salts. The alkali-metal derivatives of most of these hydrides [$NaGeH_3$, NaGe (159; *cf.* 61), $Na(K)PH_2$(133, 136; *cf.* 120), $Na(K)AsH_2$, CH_3AsHK (139), Na_4Ge_x (x = 1 or greater), M_4Sn_x, M_4Pb_x, M_3P_x, M_3As_x, M_3Sb_x, M_3Bi_x, M_2S_x, M_2Se_x, and M_2Te_x (see reference 60, for primary references)] are not ammonolyzed to any considerable extent. On the other hand, phosphonium iodide is vigorously ammonolyzed at −33°C., forming phosphine (57). (For the behavior of M_2O consult section II C and for that of MCH_3 consult section IV A.)

Anhydrous hydrazine (93) and hydroxylamine (*cf.* 4) may be readily prepared by the ammonolysis of the corresponding sulfates,

$$N_2H_4 \cdot H_2SO_4 + 2NH_3 \rightarrow N_2H_4 + (NH_4)_2SO_4$$

The ammonate of hydrazine azide, $2N_2H_5N_3 \cdot NH_3$, undergoes ammonolysis in liquid ammonia to an extent that varies directly with the temperature and with the concentration of ammonia (118).

$$2N_2H_5N_3 \cdot NH_3 + NH_3 \rightleftharpoons 2NH_4N_3 + 2N_2H_4 \quad \text{(above } -9°\text{C.)}$$

$$2N_2H_5N_3 \cdot NH_3 + 5NH_3 \rightleftharpoons 2(NH_4N_3 \cdot 2NH_3) + 2N_2H_4 \quad \text{(below } -9°\text{C.)}$$

C. OXIDES

In general, the oxides of metals are unreactive toward liquid ammonia. However, the monoxides of the alkali metals are completely ammonolyzed (for primary references *cf.* reference 21).

$$M_2O + NH_3 \rightarrow MNH_2 + MOH$$

In contrast to the monoxides, the alkali peroxides M_2O_2 and M_2O_4 are not ammonolyzed.

D. HALIDES

Observations relative to the ammonolysis of halides are assembled in table 1. Although the data are not complete, certain generalities are

TABLE 1

Ammonolysis of halides

GROUP	COMPOUNDS TREATED WITH AMMONIA	SUBSTANCES FORMED	NOTES	REFERENCES
Group IB:				
Gold	AuCl	$AuCl\cdot 12NH_3$(−28°C.)	At 20°C. $AuCl\cdot 3NH_3$ (white powder) forms	(189)
	AuBr	$AuBr\cdot 2NH_3$(+18°C.)	White powder	(189)
	AuI	$AuI\cdot 6NH_3$(−28°C.)	At 20°C. $AuI\cdot NH_3$ (white powder) forms	(189)
Group IIB:				
Beryllium......	BeX_2(X = Cl, Br, I, CN)		Probably somewhat ammonolyzed; both soluble and insoluble ammonobasic salts known	(12)
Magnesium....	$MgCl_2\cdot 2H_2O$	$MgCl_2\cdot 6NH_3$(20°C.)	$MgCl_2$ results on heating (at 290–320°C.) for 3 hr.	(192)
Mercury.......	$HgCl_2$	$Hg(NH_2)Cl$	White precipitate; ammonolysis arrested by NH_4Cl, promoted by $NaNH_2$	(78, 82; *cf.* 81)
	$HgBr_2$	Hg_2NBr	Yellow precipitate, soluble in NH_4Br; formation promoted by KNH_2	(68, 77)
	HgI_2	Hg_2NI	Yellow precipitate, soluble in NH_4I; formation promoted by KNH_2	(76)
Group IIIB:				
Boron..........	BF_3		Appreciable ammonolysis (−33°C.)	(157)
			Aminolysis in $C_2H_5NH_2$ brought to completion by Li(−33°C.)	(158)
	BCl_3	$B(NH_2)_3$(−23°, 0°C.)	Amide insoluble	(134)
	BBr_3	$B_2(NH)_3$(−10°, +20°C.)	Imide insoluble	(135)
	$B_2S_3\cdot H_2S$	$B_2S_3\cdot 6NH_3$	Yellow crystals; ammonolysis to $B_2(NH)_3$ on heating to 115–120°C.	(260)

Aluminum.....	AlI_3	$AlI_3 \cdot 20NH_3$(−33°C.) $AlI_3 \cdot 6NH_3$(8–13°C.)	Ammonolysis by KNH_2 to soluble and insoluble ammonobasic salts	(83; *cf.* 11, 80)
Gallium	$GaBr_3$	$GaBr_3 \cdot 6NH_3$(−33°C.)	No ammonolysis	(138)
	GaI_3	$GaI_3 \cdot 6NH_3$(−33°C.)	No ammonolysis	(138)
Group IIIA:				
Cerium........	CeI_3	Soluble	KNH_2 gives ammonobasic salts	(13)
Samarium......	$SmCl_3$	$SmCl_3 \cdot 13NH_3$(−78°C.)	Between −78° and 280°C., ammonates with 11.5, 8, 5, 4, $2\frac{2}{3}$, 2, and 1 NH_3 are formed	(147)
Group IVB:				
Carbon	CCl_4		No reaction at ordinary temperatures	
		$HNC(NH_2)_2$	140°C.; high pressures	(258)
Silicon.........	$SiCl_4$	$Si(NH_2)_4$(−50° to 0°C.)	Changes to $Si(NH)_2$ above 0°C.	(271)
		$HNSi(NH_2)_2$ (low temperatures)	Changes to $Si(NH)_2$(0–110°C.)	(261)
		$Si(NH)_2$(20°C.)	Benzene solution; mixture with NH_4Cl not separated	(184)
	$SiBr_4$		Ammonolysis observed (20°C.)	(91)
	$SiSCl_2$	$Si(NH)_2$		(31)
	SiS_2	$Si(NH)_2$		(31)
	SiH_3Cl	$(SiH_3)_3N$	Gaseous NH_3(20°C.); some decomposition to SiH_4 and $[SiH_2(NH)]_x$ with excess NH_3	(260a)
	SiH_2Cl_2	$[SiH_2(NH)]_x$	Gaseous NH_3(20°C.)	(260a)
	$SiHCl_3$	$[SiH(NH)]_2NH$(−110°, 20°C.)	Forms HSiN on heating to 450°C.	(260b)
	Si_2Cl_6	$[Si(NH)NH_2]_2$ (low temperatures)	Forms $Si_2N_3H_3$ at −10°C.	(237)
Germanium....	$GeCl_4$	$Ge(NH)_2$(−80°C.)	Same product in ether solution	(232, 236, 270)
			Aminolysis by $C_2H_5NH_2$, $(C_2H_5)_2NH$, $C_6H_5NH_2$, etc.	(270a)
	$GeHCl_3$	$GeHCl_3 \cdot 2NH_3$	GeNH formed on further treatment with NH_3	(51)
	GeI_4	$Ge(NH)_2$(−33°C.)	Imide obtained by extraction with liquid NH_3; same reaction in CCl_4(20°C.)	(140)
			Aminolysis by $C_2H_5NH_2$ and $(C_2H_5)_2NH$	(140)

GROUP	COMPOUNDS TREATED WITH AMMONIA	SUBSTANCES FORMED	NOTES	REFERENCES
	GeI_2	$GeNH$ (−33°C.)	NH_3 gas diluted with N_2 to avoid $2Ge^{++} \rightarrow Ge^{+4} + Ge$; GeNH obtained by extraction with NH_3	(137)
	GeS_2		Not ammonolyzed (−33°C.)	(141)
	GeS		Not ammonolyzed (−33°C.)	(141)
Tin............	$SnCl_4$	$2Sn(NH_2)_3Cl \cdot NH_4Cl$ (0°C.)	Washing with NH_3; after heating to 100°C. and again extracting $Sn(NH_2)_3Cl$ obtained	(234)
	SnI_4		Ammonobasic salt (probably) (20°C.); ammonolysis prevented by NH_4I	(70)
Lead...........	$PbCl_4 \cdot 4NH_3$		Apparently no ammonolysis (20°C.)	(57)
	$(NH_4)_2PbCl_6$		Complex ammonobasic salt by washing with NH_3	(234)
	PbI_2	$PbNH \cdot Pb(NH_2)I$ (probably)	Slight ammonolysis, prevented by NH_4I and promoted by KNH_2	(79)
Group IVA:				
Titanium......	$TiCl_4$	$TiNCl \cdot xNH_3$	Washed with NH_3	(228; *cf.* 30, 257)
	$TiCl_3$	$TiCl_3 \cdot 6NH_3$(−78°C.)	White compound	(231)
	$TiCl_2$	$TiCl_2 \cdot 4NH_3$(−78°C.)	Pearl-gray powder	(231)
	$TiBr_4$	$TiNBr \cdot xNH_3$	Yellow solid, obtained by washing with NH_3	(90, 228)
		$2Ti(NH_2)_4 \cdot TiBr_4 \cdot 8NH_3$		(230)
Zirconium.. ...	$ZrBr_4$	$3Zr(NH)_2 \cdot 7NH_4Br \cdot xNH_3$ (0–20°C.)	Ammonolysis prevented by excess NH_4Br; further ammonolysis to $Zr(NH)_2$ in presence K or KNH_2	(34, 279; *cf.* 284)
	ZrI_4		Ammonolysis indicated by extraction of NH_4I	(259)
Group VB:				
Nitrogen.......	NOCl	$NO(NH_2)$ (temperature of liquid air)	Unstable red liquid	(233)

Phosphorus ...	PCl_5	$P(NH)_2NH_2$(−50°C.)	Forms PN_2H at 400–450°C.	(199, 200)
	$PNCl_2$	$PN(NH_2)_2$(−70°C.?)	Forms PN_2H at 220°C.	(25)
	$(PNCl_2)_3$	$(PN_3H_4)_3$		(201)
	PCl_3	$P(NH_2)_3 + P_2(NH)_3$		(135)
		$P(NH_2)_3$	Forms P_4N_6 at 550°C.	(202)
	PBr_3	$P(NH_2)_3$(−70°C.)	Forms $P_2(NH)_3$ at 0°C. and PN at 250–300°C.	(122)
	PI_3	$P(NH_2)_3$(−65°C.)		(122)
Arsenic........	AsX_3(X = Cl, Br, I)	$As(NH_2)_3$(−40° to −30°C.)	Insoluble; forms $As_2(NH)_3$(0–60°C.) and AsN(250°C.)	(121)
Antimony......	$SbCl_3$	SbN	Yellow to orange solid; obtained by extraction with NH_3	(235)
Bismuth.......	$BiCl_3$	$BiCl_3 \cdot 12.5NH_3$(−78°C.)	No ammonolysis on washing; ammonates with 1(0°C.) and 0.5 NH_3	(238)
	$BiBr_3$	$BiBr_3 \cdot 18.5NH_3$(−78°C.)	Ammonates with 14(0°C.), 3, 1.5, and 0.5 NH_3	(238)
	BiI_3	$BiI_3 \cdot 22NH_3$(−78°C.)	Ammonates with 18, 16, 12.5 (0°C.), 12, 8, 4, 3, 2 NH_3	(238)
Group VA:				
Vanadium......	$VOCl_3$	$VO(NH_2)_3$(−80°C.)	Yellow solid	(213)
	VCl_3, VBr_3	$V(NH_3)_6X_3$(−33°C.)		(189a)
Tantalum......	$TaCl_5$	$Ta(NH_2)_2Cl_3 \cdot 3NH_3$	Below 0°C. ammonates of $TaCl_5$ are formed	(254; *cf.* 197–8)
Group VIIB:				
Tellurium......	$TeBr_4$	Te_3N_4(−80°C.)	Yellow solid	(268)
Group VIA:				
Chromium.....	$CrCl_3$	$[Cr(NH_3)_6]Cl_3$ + $[Cr(NH_3)_5Cl]Cl_2$	Apparently no ammonolysis	(45a)
Molybdenum ..	$MoCl_5$	$Mo(NH)_2NH_2$ or $Mo(NH_2)_4Cl$		(10)
	$MoCl_3$, $MoCl_2$		Ammonolysis and oxidation	(10)
Tungsten......	WCl_6		Ammonolysis	(10)
	WCl_5, WCl_2, WBr_2		Ammonolysis and oxidation	(10)
Uranium.......	UCl_4	$UCl_4 \cdot 12NH_3$(−78°C.)	Ammonates with 10, 6, 5, 4, and 3 NH_3	(253)
	UO_2Cl_2	$UO_2Cl_2 \cdot 10NH_3$(−78°C.)	Ammonates with 5, 4, 3, 2, and 1 NH_3	(253)

evident. Essentially, the conduct of halides (and oxyhalides) toward ammonia resembles their conduct toward water and ranges all the way from complete ammonolysis, in the case of the covalent halides, to no ammonolysis at all in the case of true salts. The principal difference is that many salts, such as those of bismuth, which are noticeably hydrolyzed, are apparently not ammonolyzed. In any particular family, ammonolysis decreases in passing from the elements of low atomic weight to those of higher atomic weight. This relationship is evident in Group III, where the boron halides are completely ammonolyzed, aluminum iodide slightly, and the gallium halides not at all; in Group IVB, where the silicon and germanium halides are completely ammonolyzed,[2] while stannic and plumbic halides form ammonobasic salts; in Group IVA, where titanium and zirconium halides form ammonobasic salts, while thorium halides appear not to be ammonolyzed; and in Group VB where phosphorus, arsenic, and antimony are completely ammonolyzed while bismuth is little if any ammonolyzed. Mercuric salts are exceptional in the ease with which they are ammonolyzed.

While samaric chloride forms ammonates, samarous chloride is oxidized by ammonia with the formation of ammonobasic samaric chloride (147).

$$2SmCl_2 + 2NH_3 \rightarrow 2Sm(NH_2)Cl_2 + H_2$$

$$Sm(NH_2)Cl_2 + 2NH_3 \rightarrow Sm(NH_2)_2Cl + NH_4Cl$$

Mono- and di-chlorogermane react with ammonia, but the reactions are not ammonolytic in character (51).

$$3xGeH_3Cl + 3xNH_3 \rightarrow 3xNH_4Cl + xGeH_4 + 2(GeH)_x \quad (-78^\circ \text{ to } -50^\circ\text{C.})$$

$$GeH_2Cl_2 + 2NH_3 \rightarrow Ge + 2NH_4Cl \quad (-27^\circ\text{C.})$$

E. SULFIDES

The behavior of sulfides toward liquid ammonia has been little investigated. Data relative to the ammonolysis of boron and silicon sulfides are included in table 1.

F. SALTS OF OXY ACIDS

Although many reactions of salts of oxygen acids have been studied in liquid ammonia, ammonolysis has been reported in only two instances, lead nitrate (79) and bismuth nitrate (89).

An ammonolytic reaction of a somewhat different type is that between various metal nitrates and the alkali amides in liquid ammonia (87).

[2] The lack of reactivity in the case of the carbon halides is unique and may be attributed to the fact that carbon is here exerting its maximum covalence, so that ammonation as a preliminary step to ammonolysis is not possible (248).

$$KNO_3 + 2KNH_2 \rightarrow KN_3 + 3KOH + NH_3$$

The above reaction takes place at 80–90°C. but better yields (up to 75 per cent), in a shorter time, result from operating at a higher temperature (130–140°C.; at 200°C. much of the azide decomposes). Other nitrates and amides likewise yield azides: sodium nitrate and sodium amide, 16 per cent; lead nitrate and excess potassium amide, 80 per cent.

G. SALTS OF AMMONO ACIDS AND AMPHOTERIC AMMONO BASES

Because the ammonolytic power of liquid ammonia is weak compared to the hydrolytic power of water, many salts of very weak acids are perfectly stable in ammonia. However, salts of extremely weak acids are ammonolyzed. Because of this tendency it is practically impossible to prepare pure specimens of potassium ammonosilicate (14, 91) and potassium ammonozirconate (34). During the washing process to remove impurities, the potassium amide formed by ammonolysis

$$Zr(NK)_2 \cdot 2NH_3 \rightarrow Zr(NH)_2 + 2KNH_2$$

is continuously removed. Ammonolysis may well be a contributing factor to the difficulty of isolating potassium ammonotitanate in pure form (90).

Evidence of the low ammonolytic power of liquid ammonia is afforded by the marked stability of salts of amphoteric bases in this solvent.

$$Zn(NH_2)_2 + 2KNH_2 \rightarrow Zn(NHK)_2 \cdot 2NH_3$$

Of the many compounds of this type (Be, Mg, Zn, Cd, Ca, Sr, Ba, La, Ce, Mo, W, Mn, Ni, Cu^{I}, Ag, Al, Ga, Tl, Sn^{II}, Sn^{IV}, and Pb^{II}; for bibliography, *cf.* references 15 and 19), only one, potassium ammonosodiate (280), has been found to be extensively ammonolyzed.

H. MISCELLANEOUS

The weak ammonolytic power of liquid ammonia is utilized to good advantage in the purification of alkali cyanides (101) and in the preparation of calcium (73) and aluminum cyanides (9).

III. Ammonolysis of Organic Compounds

A. INTRODUCTION

When one surveys the field of ammonolytic reactions in organic chemistry, one is impressed by the wide variety of substances which undergo this type of reaction. Halogen (in alkyl and aryl halides, as well as acid halides), hydroxyl (in hydroxynitriles and carbohydrates), oxygen (in aldehydes, ketones, and even certain acids and their derivatives), and sulfur may be replaced by amino or imino groups through reactions of ammonolysis. Further, Schiff bases, esters, acyl and sulfonyl derivatives,

lactones, and other similar compounds are all capable of ammonolysis. It is not difficult to envision a time in the near future when liquid ammonia reactions will have become standard practices in the organic laboratory. This will result (*1*) because liquid ammonia can bring about reactions that are not possible with aqueous or non-aqueous solutions of ammonia; (*2*) because there are inherent differences in the extent and type of reactions taking place in solutions of ammonia and in the anhydrous solvent; and (*3*) because the ease of varying conditions and thus varying both the extent and type of reaction give further variety when using liquid ammonia.

Despite the large number of investigations on the ammonolysis of organic substances in liquid ammonia, there is still need for considerable work. At present the available data are not sufficient to enable one to state definitely just what kinds of substances may be ammonolyzed, and what may not. Although some comparative studies between the action of solutions of ammonia and of the anhydrous material have been made, more information is needed to establish definitely the differences in the reactivity of the two systems. Likewise there is need for a more careful evaluation of the influence of reacting conditions (temperature, catalysis, etc.) upon ammonolysis. The catalytic effect of ammonium salts (acid catalysis) upon many ammonolytic reactions has been clearly established. More investigations in this fascinating area will be very valuable. In particular, it is desirable to establish whether or not the alkali amides exert a catalytic effect (basic catalysis) similar to that of acids.

In the discussion following, comparison is made in several cases between reactions involving liquid ammonia and those involving either aqueous or non-aqueous solutions of ammonia. For a detailed discussion of the behavior of the latter type of reactions, the reader is referred to other compilations, particularly those of Groggins (102–6).

B. HALOGEN COMPOUNDS (EXCLUSIVE OF ACYL HALIDES)

1. Alkyl halides

When the alkyl halides (chlorides, bromides, iodides) are heated to 100°C., in sealed tubes, with alcoholic ammonia, a mixture of the hydrohalides of primary, secondary, and tertiary amines, as well as the quaternary ammonium compounds results. The alkyl halides are not as reactive toward liquid ammonia as might be expected, and therefore have been used on numerous occasions for alkylations

$$\mathrm{RONa + R'X \rightarrow ROR' + NaX} \quad (66)$$

$$\mathrm{RC{\equiv}CNa + R'X \rightarrow RC{\equiv}CR' + NaX} \quad (64)$$

in liquid ammonia and for the production of olefins (18). While on certain occasions the methyl halides have been used successfully for alkylations (129, 185, 203; *cf.* 206), this practice is not to be recommended, since they are readily converted into the tetramethylammonium salts (44, 276).

$$4CH_3X + 4NH_3 \rightarrow 3NH_4X + (CH_3)_4NX$$

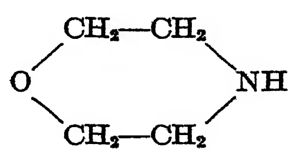
Ethyl iodide reacts very little with liquid ammonia (−33°C.), while ethyl bromide and *n*-butyl bromide do not react (276). However, at room temperature ethyl bromide and ethyl iodide react completely in 24 hr. At the end of 4 hr., there are formed 28 per cent of primary amine, some secondary amine, and a small amount of tertiary amine. Ethyl chloride under similar conditions shows no reaction in 1 hr. and very little in 24 hr. (220; *cf.* 258). Tetrakis(bromomethyl)methane, $C(CH_2Br)_4$, gives only small amounts of the corresponding amine, $C(CH_2NH_2)_4$. The same halide on heating with alcoholic ammonia (180–190°C. for 10 hr.) gives a yield of 35 per cent of the amine (100). Braun (35) has shown that *n*-amyl, *n*-octyl, and *n*-dodecyl bromides are ammonolyzed at room temperature to give a mixture of the primary and secondary amines, the yield of primary amine increasing with increasing molecular weight:

BROMIDE USED	YIELD OF PRIMARY AMINE	YIELD OF SECONDARY AMINE
	per cent	*per cent*
n-Amyl bromide	10	80
n-Octyl bromide	45	43
n-Dodecyl bromide	90	Little

On heating equal volumes of dodecyl chloride and liquid ammonia for 72–90 hr. at 75–80°C., Wibaut and coworkers (277) obtained 28 to 33 per cent of dodecylamine and 27 to 36 per cent of didodecylamine. After 170 hr. at 45°C. about 35 per cent of the chloride had reacted and, after 75 hr. at 75°C., 75 per cent had reacted.

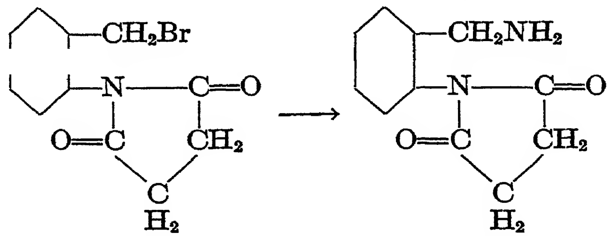
2. *Substituted alkyl halides*

Morpholine



is prepared in 80 per cent yield by treating a benzene solution of β,β'-dichlorodiethyl ether with liquid ammonia in an autoclave at 50°C. for 24

hr. (42). *N*-(α-bromo-*o*-tolyl)succinimide is ammonolyzed by liquid ammonia at room temperature over a period of 6 hr. (255).



Benzyl chloride reacts only slightly with liquid ammonia (−33°C.) (276). The ammonolysis of α-phenylethyl chloride in a sealed tube over a period

TABLE 2

Action of ammonia upon substituted alkyl halides

SUBSTITUTED ALKYL HALIDE	LIQUID AMMONIA			ALCOHOLIC AMMONIA		
	Primary	Secondary	Tertiary	Primary	Secondary	Tertiary
	per cent	*per cent*	*per cent*	*per cent*	*per cent*	*per cent*
Benzyl chloride	53	39		9	35	48
1-(Chloromethyl)naphthalene	72	20		11	38	47
9-(Chloromethyl)phenanthrene	70	26		29	25	43
β-Phenoxyethyl bromide	71			65		
N-(β-bromoethyl)-*N*-methylaniline	71	20		15	60	
N-(γ-chloropropyl)-*N*-methylaniline	65	20		18	70	
α,3,4-Trichloroquinaldine	72	22			90	
4-Anilino-α,3-dichloroquinaldine		20				
α,3-Dichloro-6-ethoxy-4-*p*-phenetidinoquinaldine	65	28		30	50	

of 6 weeks at room temperature results in the formation of the corresponding primary amine (119).

In a general examination of the action of ammonia upon substituted alkyl halides, Braun (35) has compared the yields of primary, secondary, and tertiary amines obtained by using anhydrous liquid ammonia with those using alcoholic ammonia. Three generalizations are evident: (*1*) liquid ammonia is superior to alcoholic ammonia in the preparation of primary amines; (*2*) the greater the molecular weight of the halide, the larger the percentage of primary amine produced; and (*3*) when halogen is present on the ring as well as in a side chain, only that in the side chain is replaced. The experimental results are assembled in table 2.

The simple act of treating the *N*,*N*′-bis(chloroacetyl) derivative of 1,3-propanediamine, $ClCH_2CONH(CH_2)_3NHCOCH_2Cl$, with liquid am-

monia and permitting the ammonia to evaporate yields the corresponding diamine (37). On standing for 40 hr. with liquid ammonia, the dihydrobromide of γ-bromopropylputrescine, $NH_2(CH_2)_4NH(CH_2)_3Br$, is converted to the corresponding amine, spermidine. The dihydrobromide of γ-bromopropylcadaverine on similar treatment yields *as*-homospermideine, $NH_2(CH_2)_5NH(CH_2)_3NH_2$. δ-Chlorobutylbenzamide yields small quantities of δ,δ′-dibenzamidodibutylamine, $[C_6H_5CONH(CH_2)_4]_2NH$, through a side reaction (38). Liquid ammonia rapidly acts upon ω,ω′-dibromo-*p*-ditolyl to yield the diamine (26 per cent) and a high molecular weight base, $NH_2CH_2C_6H_4C_6H_4CH_2[NHCH_2C_6H_4C_6H_4CH_2]_nNH_2$ (35).

3. Paraffin polyhalides

Whereas ethylene bromide is unreactive toward liquid ammonia (276; *cf.* 258), 1,3-propanediamine is readily prepared by simply adding 1,3-dibromopropane to liquid ammonia and allowing the latter to evaporate (37). At room temperature, 1,3-dibromopropane forms both 1,3-diaminopropane (45 to 50 per cent) and $NH_2(CH_2)_3NH(CH_2)_3NH_2$ (25 per cent). Similarly, 1,2-dichloro- or 1,2-dibromo-ethane yields ethylenediamine (65 per cent) and some diethylene triamine, while 1,11-dichloroundecane yields the diamine as the chief product. On the other hand, 1,4-dibromobutane and 1,5-dibromopentane yield bispyrrolidinium bromide and bispiperidinium bromide, respectively, as the principal products (36).

Benzotrichloride is ammonolyzed by liquid ammonia to benzonitrile in 75 per cent yield at room temperature. At 100°C., both benzonitrile and cyanphenine, $(C_6H_5CN)_3$, are formed (86).

A mixture of tri- and tetra-chloroethylenes is obtained by treating tetra- and penta-chloroethanes with an excess of anhydrous liquid or gaseous ammonia at a temperature not above 0°C. (224; *cf.* 258). 9,9-Dichlorofluorene and 1,2-dichlorodibiphenyleneëthane do not react with liquid ammonia at room temperature (221).

4. Aryl halides

While the ammonolysis (aqueous ammonia) of chlorobenzene constitutes an important industrial synthesis of aniline (107), practically no information is available on the behavior of the phenyl halides toward liquid ammonia except that iodobenzene does not react at −33°C. (276; *cf.* 258). 9-Phenyl- and 9-(α-naphthyl)-9-fluorylamines may be prepared readily by the ammonolysis of the corresponding 9-aryl-9-chlorofluorenes (at room temperature over a period of several hours). The ammonolysis of 9-methyl-9-bromofluorene is more difficult, but is accomplished by a mixture of toluene and liquid ammonia (at 75°C. for 20 hr.) (222).

5. *Heterocyclic halides*

2- and 3-aminoquinolines have been prepared in good yields by the treatment of the corresponding bromides with liquid ammonia (at 70°C. for several hours) in the presence of copper powder as a catalyst. In the absence of the copper catalyst, 2-bromoquinoline is only slowly ammonolyzed (131).

C. ALCOHOLS

Hydroxyacetonitrile is converted to aminoacetonitrile by a mixture of alcohol and liquid ammonia (1:3 by weight) (36 hr. at room temperature). Similarly, hydroxycapronitrile forms leucine nitrile with liquid ammonia (autoclave) (94; *cf.* 95). A 95 per cent yield of aminoacetonitrile is obtained by allowing a mixture of hydroxyacetonitrile and liquid ammonia to stand for 24 hr. at room temperature (188).

D. AMINES

One would expect the salt of an amine to be ammonolyzed (*1*) when the amine is less basic than ammonia or (*2*) when an insoluble compound is formed. Aniline and *p*-toluidine are liberated from their salts by simple solution in liquid ammonia (−33°C.) and evaporation of the latter. Under ideal conditions, as much as 87 per cent of the theoretical amounts of aniline may be recovered. The same treatment is partially effective with the more basic benzylamine but is ineffective with ethylamine, which is more basic than ammonia (250). Semicarbazide may be obtained in 93 per cent yield by the ammonolysis of its sulfate (4).

$$NH_2CON_2H_3 \cdot H_2SO_4 + 2NH_3 \rightarrow NH_2CON_2H_3 + (NH_4)_2SO_4$$

Liquid ammonia is not a suitable solvent for the preparation of alkali-metal salts of the aliphatic amines, since such substances are extensively ammonolyzed (225).

$$MNHR + NH_3 \rightarrow MNH_2 + H_2NR$$

Monoalkali-metal salts of aromatic amines (aniline, ethylaniline, *o*-toluidine, diphenylamine) have been prepared in liquid ammonia (142, 219, 275; *cf.* 172).

E. ALDEHYDES AND RELATED COMPOUNDS

1. *Aldehydes*

Familiar cases of ammonolysis not involving liquid ammonia are the formation of hexamethylenetetramine and hydrobenzamide from formaldehyde and benzaldehyde, respectively. The formation of aldehyde ammonia, $(CH_3CH{=}NH)_3 \cdot 3H_2O$, is essentially an ammonolysis. The

behavior of several of the higher aliphatic aldehydes (propionaldehyde, butyraldehyde, isobutyraldehyde, and heptaldehyde) toward liquid ammonia has been studied (267). The results indicate partial ammonolysis and the formation of aquo-ammono aldols.

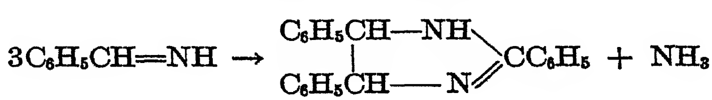
2. Acetals

The acetals are readily hydrolyzed by dilute acids but are insensitive toward bases. In contrast, the acetals are practically unaffected by liquid ammonia even over long periods of time and in the presence of ammonium salts. Dimethyl acetal yields only a trace of nitrogen-containing material after heating for 24 hr. at 130°C., either with or without ammonium chloride. Diethyl acetal gives a similar trace of nitrogenous material on standing at room temperature over a period of 3 years; it is unaffected by ammonium salts over a period of 2 years at room temperature and a period of 24 hr. at 130°C. Di-*n*-propyl and di-*n*-butyl acetals show no sign of reaction after heating for 12 hr. at 200°C. in the presence of ammonium chloride. Diethyl propional gives a trace of nitrogenous material after heating at 130°C. for 24 hr. (130).

3. Schiff bases

Anhydroformaldehyde aniline in contact with liquid ammonia (35 days) at room temperature undergoes no reaction. The addition of ammonium chloride to the solution is also without effect. On the other hand, this Schiff base is ammonolyzed to hexamethylenetetramine on heating to 150–180°C. for 12 hr. (267). Benzylideneaniline and benzylidene-*p*-toluidine are ammonolyzed (30 to 35 days) at room temperature to amarine (60 to 91 per cent) and to aniline and *p*-toluidine, respectively.

$$3C_6H_5CH{=}NC_6H_5 + 3NH_3 \rightarrow 3C_6H_5NH_2 + 3C_6H_5CH{=}NH$$

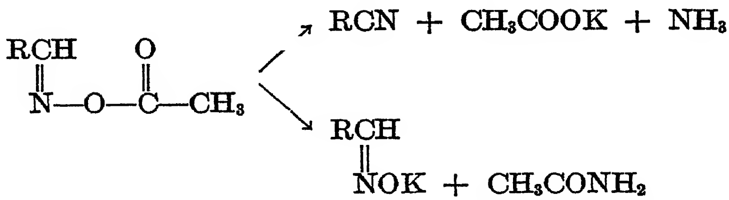


The same reactions may be accomplished in from 10 to 14 hr. by heating in the presence of ammonium chloride to 120–150°C. (264).

Benzylideneaniline, $C_6H_5CH{=}NC_6H_5$, and benzylidene-*p*-toluidine, $C_6H_5CH{=}NC_6H_4CH_3$, do not undergo simple ammonolysis with solutions of potassium amide but react in the sense of a Cannizzaro reaction (264).

$$2C_6H_5CH{=}NC_6H_5 + KNH_2 \rightarrow C_6H_5NH_2 + C_6H_5C(NCH_2C_6H_5)NKC_6H_5$$

Acetyl-α-benzaldoximes react with potassium amide in liquid ammonia (−33°C.) to form both the nitrile and the oxime:



The yields of the various products for several acetyl derivatives are assembled in table 3 (112).

Carbethoxy-α-benzaldoximes react similarly (113) (see table 4).

TABLE 3

Products obtained in the action of acetyl-α-benzaldoximes with potassium amide in liquid ammonia

ACETYL DERIVATIVE	YIELD			
	Nitrile	Acid	Oxime	Total
	per cent	*per cent*	*per cent*	*per cent*
4-Methoxy-..	47	Trace	43	90
3-Nitro-	23	60 gum	8	31
2-Chloro-.	80		9	89
Blank.		[2 gum]	[88]	[88]
3,4-Methylenedioxy	58	2	39	99
Blank		[1]	[96]	[97]

TABLE 4

Products obtained in the action of carbethoxy-α-benzaldoximes with potassium amide in liquid ammonia

CARBETHOXY DERIVATIVE	YIELD		
	Nitrile	Oxime	Total
	per cent	*per cent*	*per cent*
3-Nitro-.	60	3	63
4-Methoxy-..	70	12	82

F. KETONES AND RELATED COMPOUNDS

1. *Ketones*

After heating acetophenone with twice its volume of liquid ammonia (at 180°C. for 4 hr.), a 3 per cent yield of acetophenone imine is obtained.

$$R_2CO + NH_3 \rightleftharpoons R_2C{=}NH + H_2O$$

By repeating the experiment in the presence of a large excess of ammonobasic aluminum chloride, the yield is increased to 30 per cent. Methyl

p-tolyl ketone, fenchone, camphor, and benzophenone are ammonolyzed to the corresponding imines in the same manner (20 per cent yield in the last case). The behavior of benzil toward liquid ammonia is somewhat complex and is similar to the reaction with alcoholic ammonia. At room temperature, benzamide and imabenzil are slowly formed; on heating (at 200°C. for 2 hr.) 40 per cent of the ketone is converted into lophine, triphenylimidazole (265). The alkali and alkaline-earth amides do not promote the ammonolysis of ketones but instead form salts of the enolic forms (benzophenone forms $(C_6H_5)_2C(NH_2)ONa$ (266)). Benzophenone does not react with liquid ammonia at room temperature over a period of several weeks, whereas fluorenone gives a practically quantitative yield of fluorenone imide by the same treatment (221).

2. Ketals

Diethyl benzophenone ketal exhibits no reaction with a liquid ammonia solution of ammonium chloride over a period of 20 hr. at 120°C. (130).

3. Schiff bases

Fluorenone anil gives a 61 per cent yield of fluorenone imide when heated to 60°C. for 4 days in the presence of ammonium chloride. Heating for 20 hr. at 60°C. in the absence of the ammonium salt gives none of the imide (221).

G. CARBOHYDRATES AND THEIR DERIVATIVES

Glucose is quantitatively transformed into aminoglucose upon solution in liquid ammonia and evaporation of the solvent (204). Inulin is not ammonolyzed in liquid ammonia at −33°C. (24).

Liquid ammonia (−33°C.) dissolves the acetylated and benzoylated derivatives of any sugar compound in which the reducing group is suitably blocked (i.e., the methyl glycoside and 1,2-acetone compounds) without removing the acyl group. If, however, the acyl group is attached to the aldehyde (or ketonic) carbon atom, it is readily removed with the subsequent formation of aldehyde ammonia derivatives and amino sugars. At room temperature, liquid ammonia removes all acyl groups from the sugars (205). When cellobiose octaacetate is heated with liquid ammonia (at 55°C. for 48 hr.), 1-aminocellobiose is formed. Treatment of cellotriose with liquid ammonia likewise introduces nitrogen into the molecule (285, 286). 6-Mesyl-1,2-acetone-3,5-glucose is converted to 6-amino-1,2-acetone-3,5-glucose on standing in liquid ammonia (for 3 weeks at room temperature (116a)).

H. ACIDS AND THEIR DERIVATIVES (EXCLUSIVE OF ESTERS)

1. Acids and acid anhydrides

The preparation of acid amides by the action of liquid ammonia upon acids and acid anhydrides has been patented (209). Acetic and palmitic acids are cited as examples. The removal of water is accomplished by reacting in the presence of anhydrous salts or by evaporation with ammonia. Butyric anhydride is likewise cited as well as acetic anhydride. In the case of the latter, equimolecular quantities of ammonia and the anhydride (at 20°C. for several hours) produce acetamide and acetic acid which are separated by distillation, preferably under reduced pressure. No acetic acid is formed if two molecular proportions of ammonia are used.

When alkali bicarbonates are treated with anhydrous liquid ammonia or strong aqueous ammonia, alkali carbamates are formed (47a, 185a, 185b).

The preparation of sodium salts of amino acids by reaction of the acid with sodium in ammonia solution gives products (monosodium salts) which in the case of dicarboxylic acids, such as asparagine, are contaminated with small amounts of the monoamides (272).

2. Acid halides

Bartow and McFarland (8) first observed that acid chlorides, when dropped slowly into liquid ammonia, react vigorously with the formation of dense white fumes. Good yields of the amides of mono-, di-, and trichloroacetic and benzoic acids were obtained from the corresponding acid chlorides. Ethyl chlorocarbonate reacts vigorously, but the reaction products have not been determined. Difficulty was experienced in separating acetamide from the admixed ammonium chloride. Govaert (99) obtained acetamide in 88 per cent yield and benzamide in 98 per cent yield by adding solutions of the respective chlorides dropwise to liquid ammonia. 2-(4-Bromobenzoyl)benzoyl chloride is readily ammonolyzed (−33°C.) to 2-(4-bromobenzoyl)benzamide in 95 per cent yield (190). The ammonolysis of acid halides has been patented, and the reaction of butyryl chloride with liquid ammonia in a sealed tube has been cited as an example. Acetyl bromide and benzoyl chloride are given as further examples (209).

3. Acid amides and other mixed aquo-ammono acids

Acetamide is partially converted into acetamidine when heated with ammonium chloride in liquid ammonia (86). Urea on being heated with liquid ammonia (at 300°C. for 120 hr.) in the presence of ammonium chloride yields guanidine (19.6 per cent). Under the same conditions cyanuric

acid yields both guanidine (16.6 per cent) and urea (49.3 per cent). These ammonolytic reactions do not go to completion but result in an equilibrium mixture of guanidine, urea, and probably carbonic acid. Biuret, ammeline, ammelide, guanylurea, and triuret give, when heated (at 300°C. for 65 hr.) with ammonium chloride, considerable quantities of both urea and guanidine. Upon similar treatment, thioammeline likewise yields guanidine (26).

Of recent years there has been considerable study of the synthesis of urea from carbon dioxide and ammonia. One step in this process involves the ammonolysis of carbamic acid (ammonium carbamate) to urea. It is claimed that by operating at high pressures and temperatures as much as 79 per cent conversion may be obtained or even 81 per cent if a considerable excess (280 per cent) of ammonia over the carbamate is used (32–3, 52–4, 108, 149, 150, 183).

One might well expect that the alkali amides (ammono bases) would promote the ammonolysis of such substances as acetamide and urea. The action of alkali amides at higher temperatures has never been investigated, but at room temperatures salts are formed without any evidence of ammonolysis. Alkali-metal salts of the following substances have been prepared: acetamide, phenylacetamide, benzamide, benzenesulfonamide, toluenesulfonamide, *m*- and *p*-methoxybenzenesulfonamides (92), urea (28, 92), carbamic acid, biuret, triuret (28), and cyanourea (29).

I. ESTERS

1. Esters of carboxylic acids

By dropping the esters into a large excess of liquid ammonia (−33°C.) and allowing the mixtures to stand until the evaporation of the ammonia, Bartow and McFarland (8) observed no action with the ethyl esters of formic, acetic, propionic, valeric, caprylic, and phenylacetic acids and partial reaction with ethyl pelargonate. By the same treatment, the ethyl esters of mono-, di-, and tri-chloroacetic acids gave quantitative yields of the corresponding amides. Similarly, the ethyl esters of cyanoacetic, dibromoacetic, and chloropropionic acids gave good yields of the amides, while the esters of mono- and tri-bromoacetic acids gave unidentified reaction products (apparently the bromine atoms are partially ammonolyzed). Ethyl oxalate formed the amide in good yield, and ethyl tartrate formed it in poor yield. The ethyl esters of carbonic, malonic (*cf.* 258), succinic, lactic, levulinic, citric, benzoylacetic, benzoic, phthalic, and salicylic acids and methyl benzoate exhibited no reaction. In order to determine the effect of temperature on the ammonolysis of esters, these investigators prepared sealed tubes containing the ester and ammo-

nia (1:4 by volume). After standing 12 hr. at room temperature, ethyl acetate, pelargonate, carbonate, succinate, benzoylacetate, and benzoate had reacted no more than at −33°C., while ethyl malonate had formed malonamide. Similar standing at 60–70°C. produced no reaction with ethyl valerate, ethyl benzoylacetate, ethyl succinate, and ethyl benzoate, slight reaction with ethyl acetate, and somewhat more with ethyl tartrate. Ethyl acetoacetate (at both −33°C. and room temperature) yields with liquid ammonia a mass of white crystals rapidly changing to a yellow oil. Ethyl carbonate is ammonolyzed to both urea and guanidine when heated (at 300°C. for 65 hr.) with an ammonia solution of ammonium chloride (27).

More recently Glattfeld and Macmillan (98) have shown that butyl acetate, ethyl benzoate, methyl salicylate, and glycerol monoacetate are not ammonolyzed (−33°C.), while methyl, ethyl, propyl, and butyl lactates, ethyl mandelate, and ethyl phenylacetate are ammonolyzed (30 to 50 per cent in the case of the lactates). Ôeda (212) states that ethyl lactate and ethyl mandelate are ammonolyzed (−33°C.) to the extent of 2.5 per cent and 25 per cent, respectively (see page 32). Ethyl acetate gives a very small amount of acetamide after heating (for 24 hr. at 130°C.) with an excess of liquid ammonia (130). If the heating is carried out in the presence of ammonium chloride, acetamide (43 per cent yield) and acetamidine (small amounts) are obtained.

The experimental evidence of Shatenshtein and of Audrieth and co-workers clearly demonstrates that the ammonolysis of ethyl benzoate (0° and 25°C.) (56), diethyl malonate (−33° and 0°C.) (251), and diethyl tartrate (20°C.) (243, 245–6) is catalyzed by ammonium salts. The ammonolysis of ethyl benzoate is a pseudo first-order reaction. The catalytic effect of equivalent concentration of various ammonium salts is given by the series

$$C_6H_5COONH_4 > NH_4Cl > NH_4Br > NH_4ClO_4$$

The ammonolysis of ethyl malonate proceeds so rapidly at 0°C. that its study is not susceptible to the degree of accuracy attained in other cases. The quantitative interpretation is further complicated by the fact that ammonolysis proceeds through the intermediate malonamate stage,

$$CH_2(COOC_2H_5)_2 \rightarrow CH_2(COOC_2H_5)(CONH_2) \rightarrow CH_2(CONH_2)_2$$

However, in the early stages of the reaction the yields of malonamide were found to be proportional to the concentration of added ammonium chloride.

Audrieth and Kleinberg (5) have examined the influence of various α-substituents on the reactivity of esters toward ammonolysis in liquid ammonia (0°C.; 24 to 48 hr.). The effect of the α-substituent is given by the following series:

$$\mathrm{NC{-}C} > \mathrm{H_2NOC{-}C} > \mathrm{C_2H_5OOC{-}C} > \begin{matrix}\mathrm{C_6H_5}\diagdown \\ \mathrm{HO}\diagup\end{matrix}\!\mathrm{C} >$$

$$\begin{matrix}\mathrm{CH_3}\diagdown \\ \mathrm{HO}\diagup\end{matrix}\!\mathrm{C} > \mathrm{C_2H_5O{-}C} > \mathrm{C_6H_5{-}C} > \mathrm{H{-}C}$$

The reactivities of esters towards ammonolysis in liquid ammonia parallel, qualitatively, the reactivities of esters toward alkaline hydrolysis in aqueous solution. While the magnitude of the catalytic effect of the addition of ammonium chloride varies with different esters, there is in every case a marked increase in the yields of the corresponding acid amides. These investigators recommend liquid ammonia ammonolysis (room temperature) of esters for the preparation of α-hydroxy acids and give details for the preparation of mandelamide (81 per cent) and lactamide (71 per cent; 74 to 76 per cent with ammonium chloride).

The possibilities of utilizing ester ammonolysis commercially have not been overlooked. Many such reactions are covered by patents (208–9):

"Acid amides are prepared by reacting carboxylic acids, or their anhydrides, halides, or esters, having acid radicals containing at least two carbon atoms, with liquid anhydrous ammonia at superatmospheric pressure and room temperature. With esters, anhydrous alcohols are also obtained. The reaction is represented

$$\mathrm{R{-}\overset{\overset{\displaystyle O}{\|}}{C}{-}R_1} + \mathrm{NH_3} \rightarrow \mathrm{RCONH_2} + \mathrm{R_1H}$$

More than one =CO group may be present as in polybasic acids, e.g., malonic or succinic, or their derivatives, or polybasic alcohols partly or completely esterified with mono- and/or poly-basic acids. R is any organic radical such as methyl, propyl, butyl, amyl, phenyl, benzyl, cinnamyl, naphthyl, pyridine, and quinoline. It may contain substituents such as hydroxy, as in lactic or malic, amine groups, as in anthranilic, or halogen, in which case the amide may contain an amine group in place of halogen. R_1 may be . . . . OR, for example, in melissyl palmitate, ... Crude animal and vegetable oils, fats, waxes or resins may be used as initial materials. Catalysts for the reaction may be used. In examples: (1) coconut oil is mixed in an autoclave with excess of liquid ammonia. After standing for 12 hrs. the amides which separate are filtered off and the filtrate distilled to recover ammonia and pure glycerol. Sardine, olive, linseed, cottonseed, soya bean, or corn germ oil, or oils obtained by oxidation of paraffin wax may be similarly treated; (2) *n*-butyl acetate and excess of liquid ammonia in a sealed vessel are kept at about 20°C. by water cooling. Excess ammonia and then butyl alcohol are distilled off to give a residue of acetamide."

Oda (210, 211) has demonstrated the ammonolysis of fatty oils (olive, coconut, castor, fish, spermacetic, wood, and linseed oil) by heating them for 0.5 to 1 hr. at 100–150°C. with liquid ammonia in an autoclave.

Methyl *l*-β-hydroxybutyrate, on standing with liquid ammonia for 60 hr. at room temperature, gives a syrup from which the crystalline amide may be obtained (144). Similarly, γ-ethyl-*N*-carbobenzoxy-*d*-glutaminate (20 hr. at 15–20°C.) furnishes the ammonium salt of *N*-carbobenzoxy-*d*-glutamine (126, 207).

$$\begin{array}{c} C_2H_5O_2\cdot CH_2CH_2CHCO_2H \\ | \\ C_6H_5CH_2OCONH \end{array} \longrightarrow \begin{array}{c} H_2NCOCH_2CH_2CHCOONH_4 \\ | \\ C_6H_5CH_2OCONH \end{array}$$

2. Esters of unsaturated acids

When an ester of an unsaturated acid is treated with ammonia, addition of ammonia to the double bond as well as ammonolysis may result. Morsch (195) investigated the action of liquid ammonia at room temperature on methyl acrylate[3] and found that the following reactions take place:

$$NH_3 + CH_2{=}CHCOOCH_3 \rightarrow \underset{\text{A}}{H_2NCH_2CH_2COOCH_3}$$

$$[+ NH_3 \rightarrow H_2NCH_2CH_2CONH_2]$$

$$NH_3 + 2CH_2{=}CHCOOCH_3 \rightarrow \underset{\text{B}}{HN\begin{cases} CH_2CH_2COOCH_3 \\ CH_2CH_2COOCH_3 \end{cases}}$$

$$+ 2NH_3 \rightarrow \underset{\text{D}}{HN\begin{cases} CH_2CH_2CONH_2 \\ CH_2CH_2CONH_2 \end{cases}}$$

$$NH_3 + 3CH_2{=}CHCOOCH_3 \rightarrow \underset{\text{C}}{N\begin{cases} CH_2CH_2COOCH_3 \\ CH_2CH_2COOCH_3 \\ CH_2CH_2COOCH_3 \end{cases}}$$

$$+ 3NH_3 \rightarrow \underset{\text{E}}{N\begin{cases} CH_2CH_2CONH_2 \\ CH_2CH_2CONH_2 \\ CH_2CH_2CONH_2 \end{cases}}$$

The dependence of the yields of the various products upon the length of

[3] For the behavior of acrylic acid toward liquid ammonia see reference 124.

time for reaction is shown in table 5. The results may be compared to those obtained with alcoholic ammonia (194), as shown in table 6.

Ethyl α-ethylacrylate (200 hr.), ethyl elaidate (3 weeks), and ethyl oleate (300 hr.) are unreactive toward liquid ammonia at room temperature, while ethyl allylacetate (25 days) gives a small amount (0.05 g. from 7.5 g. of ester) of reaction product (probably an amide) (214). Ethyl β,β-dimethylacrylate (6 months) yields ethyl aminoisovalerianate but no amide (214); ethyl β,β-diethylacrylate (6 weeks) yields β,β-diethylacrylic amide (45 per cent) (215); ethyl crotonate (100 hr.) yields ethyl

TABLE 5

Effect of reaction time upon yields of products obtained by the reaction of liquid ammonia with methyl acrylate at room temperature

TIME	A	B	C	D	E
	per cent	*per cent*	*per cent*	*per cent*	*per cent*
100 hr.	2.5	42	15.5		
14 days				*ca.* 35	*ca.* 13
5 months				*ca.* 33	*ca.* 56

TABLE 6

Effect of reaction time upon yields of products obtained by the reaction of alcoholic ammonia with methyl acrylate

	TEMPERATURE	TIME	A	B	C	D	E
			per cent	*per cent*	*per cent*	*per cent*	*per cent*
Methyl alcohol:							
10 per cent NH_3	Room temperature	1 day		23	48		
20 per cent NH_3	Room temperature	1 day		2	5	26	4
Saturated	Room temperature	14 days				39.5	36.5
Saturated	Room temperature	5 months				35	44
10 per cent NH_3	100°C.	8 hr.	3	6.5			
Ethyl alcohol:							
10 per cent NH_3	100°C.	8 hr.	8	54	5		

β-aminobutyrate (55 per cent) but no amide (*cf.* 196); and methyl hydrosorbate (14 days) yields some amide but no amino ester or amino amide (214). Methyl sorbate shows very little action in 8 days, but reacts almost completely after 4 to 6 weeks to form substances which are unstable and gradually lose ammonia under vacuum. For shorter reaction periods the products are largely soluble in ether (amino esters ?) but after 3 months' action, the products are almost completely insoluble in ether (amino amides, cyclic substances ?). From the latter, up to 10 per cent of sorbamide may be isolated. The action of the ester of crotylidenemalonic

acid is similar to that of methyl sorbate, but the isolation of definite compounds has not been successful (214).

The behavior of ethyl cinnamate is recorded in table 7. Phenyl isocrotonate (2 months at room temperature) yields an oil as a primary product and crystals of phenyl isocrotonamide as a secondary product (214).

TABLE 7

Action of ethyl cinnamate with liquid ammonia

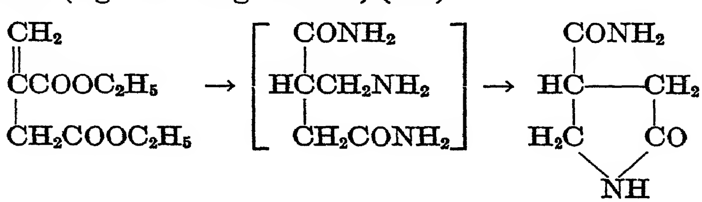
TEMPERATURE	TIME OF REACTION	UN-CHANGED ESTER	ETHYL β-AMINO-HYDRO-CINNA-MATE	CINNAMAMIDE	β-AMINOHYDRO-CINNAMAMIDE	REFER-ENCE
		per cent	*per cent*	*per cent*	*per cent*	
Room temperature	7 days	100				(218)
Room temperature	4 months			9.4	16.5	(263)
100°C.	27 hr.	55.1	22.2	14.1 per cent of amides, principally cinnamamide		(193)
100°C.	70 hr.	35	35	21.4		(193)

TABLE 8

Action of diethyl citraconate with liquid ammonia

TIME OF REACTION	UNCHANGED ESTER	DIETHYL HOMOAS-PARIGINATE	HOMOAS-PARIGIN-DIAMIDE	DIAMIDE OF CITRACONIC ACID	REFERENCES
	per cent	*per cent*	*per cent*		
1.5 days	73	5.2			(217)
5 days		Small amount			(262)
6 days....		Largely		Small amount	(217)
1 month.............			Some		(262)
2 months.		10.5	28.3		(262)
3.5 months		1.4	73		(262)
7 months		14.1			(262)
Bomb; 12 days.			51.4	Some	(262)

Diethyl itaconate (6 days) yields a hygroscopic product from which only a small quantity of itaconamide has been isolated (216). On longer standing (2 to 3 months), diethyl itaconate yields the amide of 2-oxypyrrolidine-4-acid (7 g. from 14 g. of ester) (262).

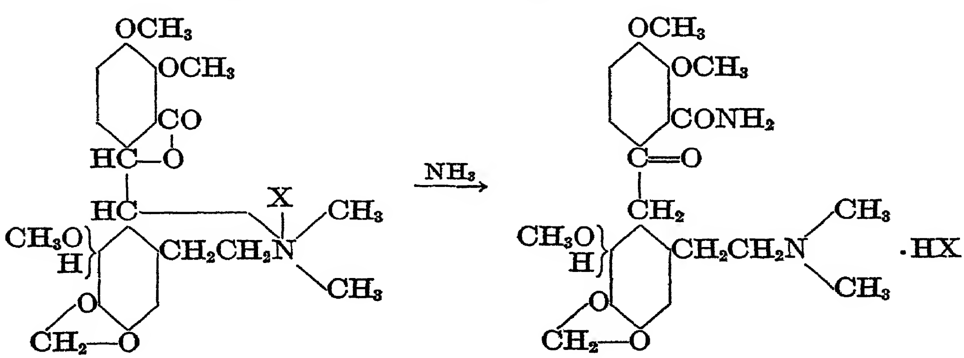


α-Mesaconic monoethyl ester (1 month) yields the corresponding amide and homoasparagine (0.56 g. from 4 g. of the acid ester). Diethyl mesaconate (3 weeks) yields mesacondiamide (0.8 g. from 9 g. of ester) (262). The results obtained with diethyl citraconate are assembled in table 8.

The ammonolysis of β-diethylaminoethyl β-chlorocrotonate, $CH_3CCl{=}CHCO_2(CH_2)_2N(C_2H_5)_2$, is interesting, since the chlorine is quantitatively removed (room temperature for 72 hr.) without ammonolysis of the ester or addition to the double bond (247).

3. Lactones

As might be expected from their relation to esters, many lactones are ammonolyzed in liquid ammonia by merely dissolving the corresponding lactones in an open vessel containing ammonia. Glattfeld and Macmillan (97; *cf.* 132) prepared the following amides in quantitative yield: *dl*-1,3-dihydroxybutyramide, *dl*-2,3-dihydroxybutyramide, *dl*-erythronamide, *d*-erythronamide, *l*-erythronamide, *d*-galactonamide, *d*-gluconamide, and *d*-mannonamide (from both the γ- and δ-lactones). Coumarin and the lactone of γ-hydroxybutyric acid are not affected by liquid ammonia at its boiling point, although the latter is ammonolyzed at room temperature in a sealed tube. Later work (98) confirmed the results on *d*-glucono-γ-lactone and coumarin and demonstrated the complete ammonolysis of 3-benzalphthalide (amide of desoxybenzoincarbonic acid). Phthalide, 3-phenylphthalide, phthalophenone, and the lactone of 2′,4′-dihydroxydiphenyl-2-carboxylic acid are not ammonolyzed (−33°C.). Phenolphthalein is temporarily ammonolyzed, but the product loses its ammonia completely on prolonged evacuation. γ-Butyrolactone and γ-valerolactone on heating for 2 to 3 hr. at 200–230°C. with liquid ammonia are converted into α-pyrrolidone (64 per cent) and 5-methyl-2-pyrrolidone (74 per cent), respectively (256). Simple solution of 2,4,6-trimethyl-

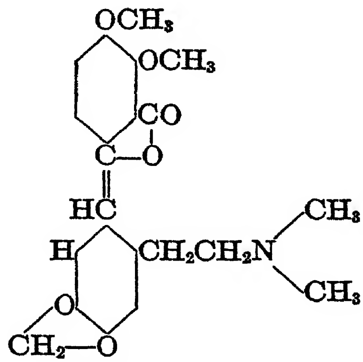


$\left.\begin{matrix}CH_3O\\H\end{matrix}\right\}$: CH_3O = narcotine methiodide; H = hydrastine methiodide

δ-gluconolactone and 3,4,6-trimethyl-δ-altronolactone in liquid ammonia (−33°C.) yields the corresponding amides (182a).

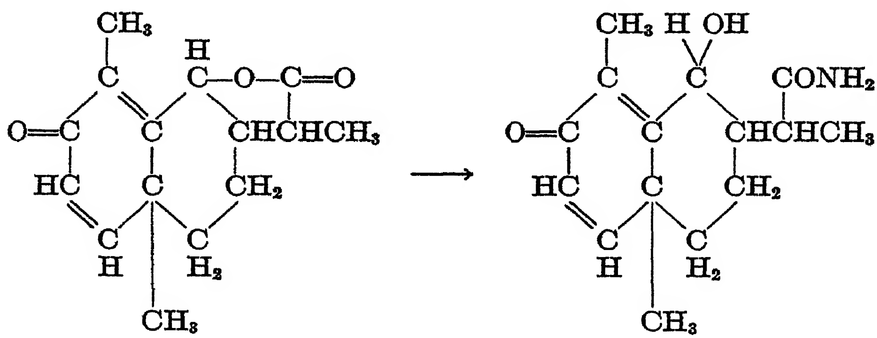
Narcotine methiodide and methochloride are converted to narceinamide hydroiodide and hydrochloride, respectively, by the action of liquid ammonia (−80° to −33°C.)

Under the same conditions hydrastine methiodide yields methylhydrastamide hydroiodide, while methylhydrastine and its hydroiodide are unaffected (2).



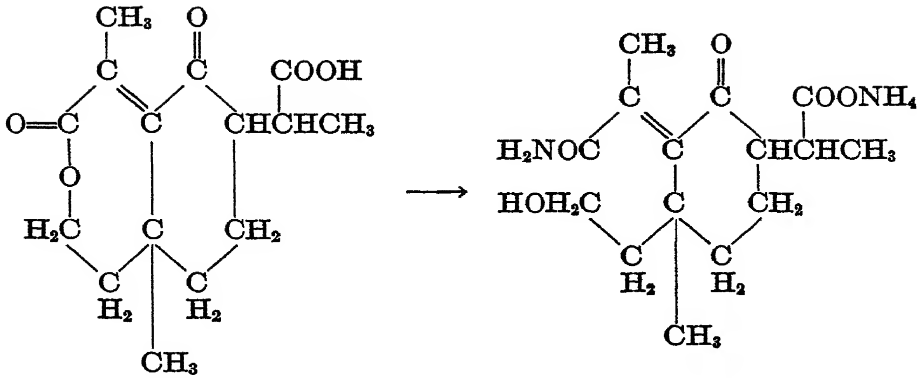
Methylhydrastine

Santonin is quantitatively ammonolyzed to the amide of santoninic acid (room temperature; 3 days) (1).

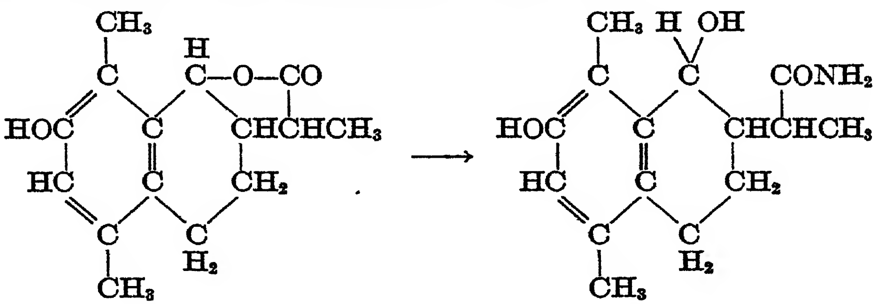


The lactone obtained by the oxidation of santonic acid is likewise ammonolyzed by liquid ammonia.

By following the course of the reaction polarimetrically, Shatenshtein has clearly demonstrated that the ammonolysis of santonin is catalyzed by ammonium salts, as well as by acid amides and other weak acids (239–42, 246). The order of catalytic activity for "strong" acids varies with the anion as follows: $ClO_4^- < I^- < NO_3^- < Br^- < Cl^-$, which is the reverse of the order of osmotic coefficients and of coefficients of electrical conduc-

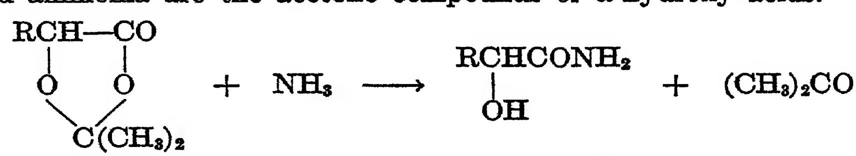


tivity of solutions of ammonium salts in liquid ammonia. The order of catalytic activity of the amides of carboxylic acids in liquid ammonia corresponds to that of the acids themselves in water ($CH_3COOH < C_6H_5COOH < HCOOH$). Such studies have also been extended to the ammonolysis of desmotroposantonin (243–4, 246).



The available evidence confirms Glattfeld's statement (98) that in contrast to simple esters lactones are never partially ammonolyzed. They are either completely ammonolyzed or not ammonolyzed at all.

Compounds closely resembling the lactones in their ease of reaction with liquid ammonia are the acetone compounds of α-hydroxy acids.



The results of Ôeda's studies are assembled in table 9 (212).

4. Esters of ortho acids

Ethyl orthocarbonate is very unreactive toward liquid ammonia. No reaction occurs on heating for 1.5 hr. at 100°C. both with and without

ammonium chloride. Traces of guanidine can be identified after 3 years (at room temperature) both with and without ammonium chloride or after heating for 24 hr. at 130°C. in the presence of ammonium chloride. Ethyl orthoformate gives no indication of reaction when heated with liquid ammonia (at 100°C. for 12 hr.) or with a solution of ammonium chloride (at 200°C. for 9 hr.) (130). This lack of reactivity of ortho esters stands in sharp contrast to the preparation of guanidine by heating esters of orthocarbonic acid with aqueous ammonia.

5. Esters of carboxazylic and mixed aquo-ammonocarbonic acids

(*a*) *Substituted acyl amides.* A convenient way to deacetylate 1-acetyl-1,2,3-benzotriazole is that of simply dissolving the material in liquid ammonia (−33°C.), allowing the ammonia to evaporate, and working over the resulting mixture (43). 1-Benzoyl-1,2,3-benzotriazole is debenzoylated in the same manner.

TABLE 9

Action of acetone compounds of α-hydroxy acids with liquid ammonia

STARTING MATERIAL	CONDITIONS	PRODUCT			
		Lactamide	Leucic acid amide	Mandelamide	Phenylacetamide
		per cent	*per cent*	*per cent*	*per cent*
Acetone compound	100°C.; 5 hr.	47	49	70	55
	Room temperature; 2 weeks		79	92	88
	ca. −33°C.; overnight	85	90	95	90
Ethyl ester	*ca.* −33°C.; overnight	<2.5		25	

Acetanilide exhibits no reaction with liquid ammonia (24 hr. at 130°C.), but in the presence of ammonium chloride gives a quantitative yield of aniline (130) and some acetamidine (86).

Ammonia solutions of potassium amide at room temperature cause no ammonolysis of substituted acyl amides, benzylacetamide and *p*-phenetolacetamide, but instead form salts (84).

(*b*) *Esters of aquo-ammonocarbonic acids.* Although ethyl allophanate is completely ammonolyzed by aqueous ammonia (at 100°C.) (117), anhydrous liquid ammonia (or ammonia containing 2 per cent of water) is without effect (48). Methyl and ethyl allophanate, urethan, carbethoxycyanamide, *s*-dicarbethoxyguanidine, methylurea (*cf.* 86), *s*-dicarbethoxyurea, and carbethoxy-*N*-phenylbiuret when heated (at 300°C. for 65 hr.) with ammonia solutions of ammonium chloride yield guanidine and urea (27). Carbanilide on similar treatment is ammonolyzed to aniline (77 per cent), urea, and some guanidine (86). With ammonia alone

(at 130°C. for 24 hr.), aniline (89 per cent) and urea are formed but no guanidine (130). Diarylguanidines are prepared by treatment of the corresponding thioureas with liquid ammonia (temperatures up to 25°C.) in the presence of sulfur-binding metal compounds, such as lead oxide (249).

$$(o\text{-}CH_3C_6H_4NH)_2CS + 3NH_3 \rightarrow (o\text{-}CH_3C_6H_4NH)_2C{=}NH + (NH_4)_2S$$

Ethyl carbamate reacts with an ammonia solution of potassium amide (at −33°C. or room temperature) to form potassium cyanate and ethyl alcohol.

$$H_2NCOOC_2H_5 + KNH_2 \rightarrow KNCO + C_2H_5OH + NH_3$$

This reaction constitutes a clear case of ammonolysis being brought about by an ammono base. With potassium amide allophanic acid forms a salt which does not decompose below 215°C. (28).

$$KHNCONHCOOC_2H_5 \rightarrow KNCO + H_2NCOOC_2H_5$$

6. Esters of carbazylic and ammonocarbonic acids

(*a*) *Esters of carbazylic acids.* N,N'-diphenylacetamidine and N,N'-diphenylbenzamidine on heating with ammonia solutions of ammonium chloride yield aniline and acetamidine and benzamidine, respectively (86). N,N'-diphenylacetamidine with ammonia alone (at 130°C. for 24 hr.) is 13 per cent ammonolyzed (130).

(*b*) *Esters of ammonocarbonic and ammonocarbonous acids.* N,N'-diphenylguanidine is partially ammonolyzed (at 130°C. for 24 hr.) to yield aniline (18 per cent) and guanidine. With identical treatment, N,N',N''-triphenylguanidine is almost completely ammonolyzed to N,N'-diphenylguanidine (130). However, heating (200°C.) triphenylguanidine with a solution of ammonium chloride yields guanidine and aniline (82 per cent) (86).

Triphenylguanidine is not ammonolyzed by ammonia solutions of potassium and sodium amides (at room temperature) but forms salts (85).

When ethyl isocyanide and an equivalent amount of potassium amide are heated (at 80°C. for a few hours) in liquid ammonia solution, ammonolysis results in accordance with the following equation (85a):

$$C_2H_5NC + KNH_2 \rightarrow C_2H_5NH_2 + KCN$$

7. Esters of sulfur and selenium acids

The alkyl esters of sulfuric acid and the sulfonic acids are attacked little if at all by liquid ammonia at low temperatures, since these substances have been used successfully for alkylations in ammonia: methyl, ethyl (115–6, 123, 145, 187), *n*-propyl, isopropyl, *n*-butyl, and *n*-amyl sulfates;

n-propyl and *n*-butyl *p*-toluenesulfonates (148). On standing for 24 hr. at room temperature with an excess of liquid ammonia, the tribenzenesulfonate of pyrogallol is one-third ammonolyzed to the dibenzenesulfonate (273).

The reaction of halogen sulfonic acid esters with liquid ammonia produces a mixture of amines. Thus, chlorosulfonic acid dodecyl ester dissolved in ether reacts at $-30°C.$ to produce a mixture of the corresponding primary, secondary, and tertiary amines.

$$ROSO_2Cl + 3NH_3 \rightarrow RNH_2 + HOSO_2NH_2 + NH_4Cl$$

(R = octyl, myristyl, cetyl, octodecyl, oleyl, etc.; olefinic radical such as that of olein alcohol; polyethyleneglycol monoalkyl ethers; aliphatic radicals of low molecular weight containing aliphatic radicals of high molecular weight, etc.) (125).

N,N'-dimethyl-*N,N'*-dinitromethionamide reacts instantly with liquid ammonia to form methionamide and methylnitramine (6).

$$CH_2[SO_2N(NO_2)CH_3]_2 + 2NH_3 \rightarrow 2CH_3NHNO_2 + CH_2[SO_2NH_2]_2$$

Methyl and ethyl selenites are ammonolyzed (at room temperature) but the expected aquo-ammonoselenite is not obtained. Instead, a mixture of selenium nitride, selenium dioxide, and selenium results (269).

8. Nitrosamines

Di-*p*-tolyl nitrosamine exhibits no reaction with a liquid ammonia solution of ammonium chloride (at room temperature for 10 days) (274). Heated with ammonia (for 24 hr. at 130°C.), di-*p*-tolyl nitrosamine gives a 51 per cent yield of nitrogen according to the reaction

$$R_2NNO + NH_3 \rightarrow R_2NH + H_2O + N_2$$

Similar treatment in the presence of ammonium chloride increases the yield of nitrogen to 61 per cent (130). Alkali amides, on the other hand, (0–20°C.) produce complete ammonolysis in the sense of the above reaction (57a).

J. MISCELLANEOUS

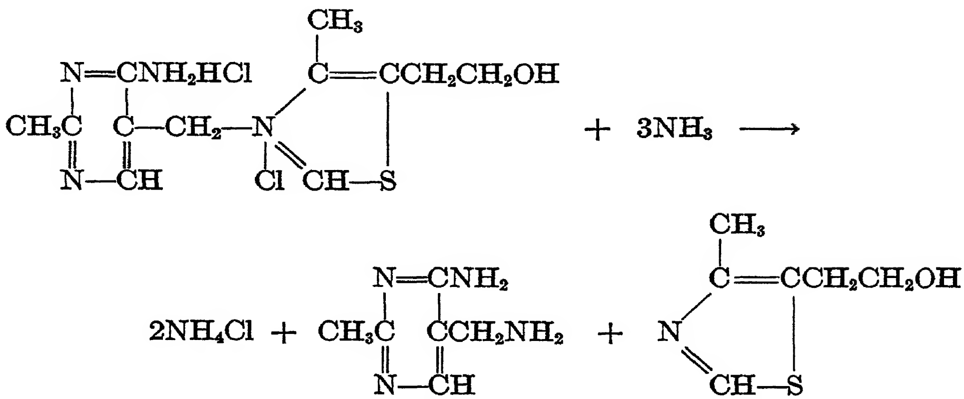
Two of the hydroxyl groups in gossypol are ammonolyzed (at $-33°C.$) to form diaminogossypol (191).

$$C_{30}H_{24}O_2(OH)_6 + 2NH_3 \rightarrow C_{26}H_{24}O_2(OH)_4(>\overset{|}{C}{=}C-NH_2)_2 + 2H_2O$$

The same compound is formed by the action of ammonia on anhydrogossypol,

$$C_{24}H_{20}O_2(OH)_2(>\overset{|}{C}{=}C{-}O{-}C{\equiv})_2$$

Thiamine or vitamin B_1 is split by liquid ammonia at room temperature into two fragments (46, 278):



Proteins are partially ammonolyzed by ammono bases (slowly by potassium amide at −33°C., not appreciably by sodium amide below 40°C.) and by ammonium salts (at 35–115°C. for 2 days) in liquid ammonia (186, 226–7).

IV. Ammonolysis of Organometallic Compounds

A. METAL ALKYLS AND ARYLS

There is no record at hand of any investigations concerning the action of liquid ammonia upon alkali alkyls, aryls, or simple aryl-substituted alkyls. However, the many reactions in ammonia solution by which these materials might well be expected to be formed, always yield an alkali amide and hydrocarbon (*cf.* 20, 62).

$$R^- + NH_3 \rightarrow RH + NH_2^-$$

Additional support to this view is given by the fact that lithium salts of the primary and secondary aliphatic and secondary aliphatic–aromatic amines are readily prepared by treating the amine with a solution of lithium phenyl or lithium *n*-butyl in absolute ether (*cf.* 22).

$$R_2NH + LiC_6H_5 \rightarrow R_2NLi + C_6H_6$$

The polyphenyl-substituted metal alkyls are less ammonolyzed than the unsubstituted ones. In general, extensive ammonolysis does not take place (at −33°C.) when an alkali metal, M, is present in the grouping $(C_6H_5)_2CM$— (17, 63, 281–2). Some indication of the influence of the

metal upon the extent of ammonolysis is realized among the salts of triphenylmethane. Potassium triphenylmethide is stable in liquid ammonia and the sodium salt is slightly ammonolyzed, while the calcium salt is completely ammonolyzed (173a). On the other hand, alkali salts of acetylene and monosubstituted acetylenes are entirely stable in ammonia and serve as useful synthetic reagents (16, 63).

1,2,3,4-Tetrasodium-1,2,3,4-tetrahydronaphthalene is extensively ammonolyzed (283).

$$C_{10}H_8Na_4 + 4NH_3 \rightleftharpoons C_{10}H_{12} + 4NaNH_2$$

At −33°C. this reaction is three-fourths complete and near room temperature is entirely complete. The above equilibrium suggests the probable existence of similar alkali derivatives of other polycyclic hydrocarbons at low temperatures (*cf.* 65). 9,10-Disodium- and 9,10-dipotassium-9,10-diphenyl-9,10-dihydroanthracenes are stable in liquid ammonia (128).

Zinc ethyl reacts completely with ammonia,

$$Zn(C_2H_5)_2 + 2NH_3 \rightarrow Zn(NH_2)_2 + 2C_2H_6$$

either in the gaseous state (69, 74) or in solution in anhydrous ether (143). A similar reaction of beryllium alkyls should prove useful in preparing the at present unknown beryllium amide. Mercury dialkyls are not ammonolyzed (151, 174). Similarly, the fully alkylated and arylated compounds of the elements in the third, fourth, fifth, and sixth groups of the periodic classification appear to be unreactive toward ammonia (except for ammonation of alkyls of boron, etc., giving $R_3B \cdot NH_3$). On the other hand, triethylsilane is ammonolyzed in liquid ammonia and aminolyzed in ethylamine (176),

$$2(C_2H_5)_3SiH + NH_3 \rightarrow [(C_2H_5)_3Si]_2NH + 2H_2$$

$$(C_2H_5)_3SiH + C_2H_5NH_2 \rightarrow (C_2H_5)_3SiNHC_2H_5 + H_2$$

The first reaction takes place in the presence of potassium amide. Lithium metal serves as a catalyst for the second reaction.

A number of alkali-metal derivatives of alkyl- and aryl-substituted silanes, germanes, and stannanes have been prepared. These include $(C_2H_5)_3SiLi$ (176), $(C_6H_5)_3SiNa(Li)$ (160, 173), $(C_2H_5)_3GeK(Li)$ (162), $(C_6H_5)_3GeNa(Li)$ (166, 173), $(C_6H_5)_2GeNa_2$ (155, 166), $NaGe(C_6H_5)_2 \cdot Ge(C_6H_5)_2Na$ (155), $(CH_3)_3SnNa$ (160, 168, 175, 178), $(CH_3)_2SnNa_2$ (41, 45, 170), $NaSn(CH_3)_2 \cdot Sn(CH_3)_2Na$ (170, 175), $(C_6H_5)_3SnNa$ (45, 173), $(C_6H_5)_2SnNa_2$ (153), $NaSn(C_6H_5)_2 \cdot Sn(C_6H_5)_2Na$ (153), and $(C_6H_5)_3PbNa$ (72a).

There is no indication of ammonolysis of any of these compounds except

lithium triethylgermanide, which is partially aminolyzed in ethylamine and completely ammonolyzed in liquid ammonia, and disodium diphenylstannide which is ammonolyzed at high dilutions. The stability of this type of salt makes possible their extensive use in synthetic work.

B. ORGANOMETALLIC HALIDES

Grignard reagents react with ammonia in much the same way that they react with water (47, 252).

$$(C_6H_5)_2C{=}NMgX + NH_3 \rightarrow (C_6H_5)_2C{=}NH + MgX(NH_2)$$

The destruction of the Grignard reagent by means of ammonia is particularly advantageous where, as in the above case, water would destroy the ketimine.

Kraus and coworkers have shown that the organomercuric halides, RHgX ($R = CH_3$ to C_8H_{17}, C_6H_5) (151, 169, 174), and dimethylgallium chloride, $(CH_3)_2GaCl$, are not ammonolyzed by the action of liquid ammonia (181).

In contrast, it has been found that triarylsilicyl halides, $(C_6H_5)_3SiCl$ and $(C_6H_5)_3SiBr$, are completely ammonolyzed (161, 163, 177), while the ammonolysis of triethylsilicyl bromide, $(C_2H_5)_3SiBr$, is brought to completion through the aid of metallic lithium (176). On the other hand, the trialkylsilicyl fluorides (ethyl, propyl, butyl, and amyl) are very unreactive toward liquid ammonia and even resist the action of sodium in liquid ammonia (96).

Like organosilicon halides, organogermanium halides are ammonolyzed. However, temperature seems to play an important part in the ammonolysis of triethylgermanyl bromide. At −33°C. this compound forms an ammonate, while at room temperature ammonolysis is marked and may be brought to completion by the use of sodium (164, 180). Triphenylgermanyl fluoride is also ammonolyzed (167) but not as readily as the triphenylgermanyl bromide (182), whose reaction is brought to completion by the use of potassium amide. The resulting product, triphenylgermanyl amide, very easily loses ammonia with the formation of the nitride. This property of losing ammonia is also encountered with the ethylgermanium tribromide and triiodide, in that they do not form the amide but pass directly to ethylgermanium nitride (72).

From the reaction of dialkyl- and diaryl-germanium dihalides, $(C_2H_5)_2GeBr_2$ (71) and $(C_6H_5)_2GeCl_2$ (154), one obtains the corresponding imides. It has also been shown that the organosilicon and organogermanium halides do not undergo aminolysis with ethylamine: $(C_6H_5)_3SiBr$ (161); $(C_2H_5)_3GeCl$, $(C_2H_5)_3GeBr$ (164); $(C_6H_5)_2GeCl_2$ (156).

As one passes from the organogermanic halogen compounds, which are

ammonated at $-33°C.$ but ammonolyzed at higher temperatures, one finds that the organostannic halides form ammonated products but no ammonolytic products even at room temperature: $(CH_3)_3SnCl$ (169); $(CH_3)_3SnBr$ (179); $(CH_3)_3SnI$ (169); $(CH_3)_2SnBr_2$ (171). Triphenyllead chloride and iodide are not ammonolyzed in liquid ammonia (72a, 72b).

Hein and coworkers were able to electrolyze tetraphenylchromium iodide in liquid ammonia, thereby obtaining tetraphenylchromium. It is thus evident that these compounds are relatively stable in liquid ammonia (114).

V. References

(1) Abkin, A., and Medvedev, S.: J. Gen. Chem. (U.S.S.R.) **4**, 1407–14 (1934); *cf.* Chem. Abstracts **29**, 3682 (1935).
Adams, R.: See Miller, R. F.
(2) Addinall, C. R., and Major, R. T.: J. Am. Chem. Soc. **55**, 2159–61 (1933).
(3) Akhumov, E. I., and Vasil'ev, B. B.: Russian patent 43,634; Chem. Abstracts **31**, 7611 (1937).
Anderson, E. G. E.: See White, G. F.
(4) Audrieth, L. F.: J. Am. Chem. Soc. **52**, 1250–1 (1930).
Audrieth, L. F.: See also Fellinger, L. L., and Slobutsky, C.
(5) Audrieth, L. F., and Kleinberg, J.: J. Org. Chem. **3**, 312–16 (1938).
Bachman, G. B.: See Miller, H. F.
(6) Backer, H. J.: Rec. trav. chim. **47**, 942–9 (1928).
(7) Bankhardt, A. B.: Thesis, The Ohio State University, 1933.
(8) Bartow, E., and McFarland, D. F.: Kansas Univ. Quart. **10**, 79–85 (1901); *cf.* Review Am. Chem. Research **8**, 303–4 (1902).
(9) Bergstrom, F. W.: J. Am. Chem. Soc. **46**, 1559–68 (1924).
(10) Bergstrom, F. W.: J. Am. Chem. Soc. **47**, 2317–25 (1925).
(11) Bergstrom, F. W.: J. Phys. Chem. **29**, 160 (1925).
(12) Bergstrom, F. W.: J. Am. Chem. Soc. **50**, 657–62 (1928).
(13) Bergstrom, F. W.: J. Am. Chem. Soc. **59**, 1374–5 (1937).
(14) Bergstrom, F. W.: Private communication.
Bergstrom, F. W.: See also Smith, G. E. P., Jr.
(15) Bergstrom, F. W., and Fernelius, W. C.: Chem. Rev. **12**, 51–2 (1933).
(16) Reference 15, pp. 87–90.
(17) Reference 15, pp. 90–1.
(18) Reference 15, pp. 92–3.
(19) Bergstrom, F. W., and Fernelius, W. C.: Chem. Rev. **20**, 416 (1937).
(20) Reference 19, pp. 417–18.
(21) Reference 19, p. 419.
(22) Reference 19, p. 429.
(23) Reference 19, pp. 431–2.
(24) Bergstrom, F. W., and Gilmore, A. E.: J. Am. Chem. Soc. **59**, 1356–8 (1937).
(25) Besson and Rosset: Compt. rend. **146**, 1149–51 (1908).
Bien, P. B.: See Kraus, C. A.
(26) Blair, J. S.: J. Am. Chem. Soc. **48**, 87–95 (1926).
(27) Reference 26, p. 94.
(28) Blair, J. S.: J. Am. Chem. Soc. **48**, 96–103 (1926).
(29) Blair, J. S., and Smith, G. E. P., Jr.: J. Am. Chem. Soc. **56**, 907–10 (1934).
Blix, M.: See Stock, A.

(30) Blix, M., and Wirbelauer, W.: Ber. **36,** 4220 (1903).
(31) Reference 30, pp. 4224–8.
(32) Bolotov, B. A., and Popova, A. N.: J. Chem. Ind. (Moscow) **1934,** No. 9, 32–8; *cf.* Chem. Abstracts **29,** 457 (1935).
(33) Bolotov, B. A., Popova, A. N., and Sokolova, Y. K.: J. Chem. Ind. (U.S.S.R.) **14,** 631–6 (1937); *cf.* Chem. Abstracts **31,** 6619 (1937).
(34) Bowerman, E. W., and Fernelius, W. C.: J. Am. Chem. Soc. **61,** 121–4 (1939).
(35) Braun, J. von, and others: Ber. **70B,** 979–93 (1937).
(36) Reference 35, pp. 991–3.
(37) Braun, J. von, Dengel, F., and Jacob, A.: Ber. **70B,** 996–7 (1937).
(38) Braun, J. von, and Pinkernelle, W.: Ber. **70B,** 1230–40 (1937).
Brodkorb, F.: See Gesellschaft für Kohlentechnik.
(39) Brönsted, J. N., and Grove, C.: J. Am. Chem. Soc. **52,** 1394–1403 (1930).
(40) Brönsted, J. N., and Wynne-Jones, W. F. K.: Trans. Faraday Soc. **25,** 59–76 (1929).
Brown, C. L.: See Kraus, C. A.
Brown, E. H.: See Kraus, C. A.
Browne, A. W.: See Howard, D. H., Jr.
(41) Bullard, R. H., and Robinson, W. B.: J. Am. Chem. Soc. **49,** 1368–73 (1927).
Cady, H. P.: See Franklin, E. C.
(42) Campbell, A. W.: U. S. patent 2,034,427; Chem. Abstracts **30,** 2986 (1936).
Campbell, R. W.: See Slobutsky, C.
(43) Cappel, N. O., and Fernelius, W. C.: Unpublished observations.
Carney, E. S.: See Kraus, C. A.
(44) Chablay, E.: Ann. chim. [9] **1,** 478–9 (1914).
(45) Chambers, R. W., and Scherer, P. C.: J. Am. Chem. Soc. **48,** 1054–62 (1926).
Chase, E. F.: See Kilpatrick, M., Jr.
(45a) Christensen, O. T.: Z. anorg. Chem. **4,** 229–31 (1893).
Clark, K. G.: See Krase, H. J.
(46) Cline, J. K., Williams, R. R., Ruehle, A. E., and Waterman, R. E.: J. Am. Chem. Soc. **59,** 530–3 (1937).
(47) Cornell, E. F.: J. Am. Chem. Soc. **50,** 3314–18 (1928).
(47a) Cunningham, G. L., MacMullin, R. B., and McCready, W. K. (to The Mathieson Alkali Works): British patent 409,871; Chem. Abstracts **28,** 6159 (1934).
(48) Dains, F. B., and Wertheim, E.: J. Am. Chem. Soc. **42,** 2304 (1920).
Dansetti, A.: See Pascal, P.
Davis, H. M.: See Heisig, G. B.
(49) Dawson, H. M., and Lowson, W.: J. Chem. Soc. **1927,** 2107–14.
(50) Dawson, H. M., and Lowson, W.: J. Chem. Soc. **1928,** 2152.
Dengel, F.: See Braun, J. von.
Denk, B.: See Stähler, A.
(51) Dennis, L. M., and Work, R. W.: J. Am. Chem. Soc. **55,** 4486–9 (1933).
(52) DeRopp, H. W. (to E. I. du Pont de Nemours and Co.): U. S. patent 2,083,010; Chem. Abstracts **31,** 5385 (1937).
(53) DeRopp, H. W., and Hetherington, H. C. (to The Canadian Industries, Ltd.): Canadian patent 343,938; Chem. Abstracts **28,** 7268 (1934).
(54) DeRopp, H. W., and Hetherington, H. C. (to E. I. du Pont de Nemours and Co.): U. S. patent 2,017,588; Chem. Abstracts **29,** 8220 (1935).
Dix, W. M.: See Foster, L. S.
du Pont, E. I., de Nemours and Co.: See DeRopp, H. W.
Eatough, H.: See Kraus, C. A.

EDWARDS, W. R., JR.: See Singleton, H. M.
EISNER, F.: See Ruff, O.
FASOLD, K.: See Juza, R.
(56) FELLINGER, L. L., AND AUDRIETH, L. F.: J. Am. Chem. Soc. **60**, 579–81 (1938).
(57) FERNELIUS, W. C.: Unpublished observations.
FERNELIUS, W. C.: See also Bergstrom, F. W., Bowerman, E. W., and Cappel, N. O.
(57a) FERNELIUS, W. C., AND WATT, G. W.: J. Am. Chem. Soc. **55**, 3482–5 (1933).
(58) FERNELIUS, W. C., AND WATT, G. W.: Chem. Rev. **20**, 202–4 (1937).
(59) Reference 58, pp. 206–17.
(60) Reference 58, pp. 207–15.
(61) Reference 58, pp. 212–13.
(62) Reference 58, pp. 217–22.
(63) Reference 58, pp. 220–2.
(64) Reference 58, p. 222.
(65) Reference 58, p. 223.
(66) Reference 58, p. 235.
(67) FITZGERALD, E., AND LAPWORTH, A.: J. Am. Chem. Soc. **93**, 2163–75 (1908).
(68) FITZGERALD, F. F.: J. Am. Chem. Soc. **29**, 659–60 (1907).
(69) Reference 68, pp. 663–4.
(70) FITZGERALD, F. F.: J. Am. Chem. Soc. **29**, 1695 (1907).
(71) FLOOD, E. A.: J. Am. Chem. Soc. **54**, 1667 (1932).
(72) FLOOD, E. A.: J. Am. Chem. Soc. **55**, 4935–8 (1933).
FLOOD, E. A.: See also Kraus, C. A.
FLUCK, L. A.: See Foster, L. S.
FOSTER, L. S.: See Kraus, C. A.
(72a) FOSTER, L. S., DIX, W. M., AND GRUNTFEST, I. J.: J. Am. Chem. Soc. **61**, 1685–7 (1939).
(72b) FOSTER, L. S., GRUNTFEST, I. J., AND FLUCK, L. A.: J. Am. Chem. Soc. **61**, 1687–90, (1939).
(73) FRANK, H. H., AND FREITAG, C.: Z. angew. Chem. **39**, 1430–2 (1926).
(74) FRANKLAND, E.: Phil. Mag. [4] **15**, 149–52 (1858).
(75) FRANKLIN, E. C.: J. Am. Chem. Soc. **27**, 826–7 (1905).
(76) Reference 75, pp. 838–9.
(77) Reference 75, pp. 839–41.
(78) Reference 75, pp. 841–2.
(79) Reference 75, pp. 842–7.
(80) Reference 75, p. 849.
(81) FRANKLIN, E. C.: J. Am. Chem. Soc. **29**, 35–66 (1907).
(82) FRANKLIN, E. C.: Am. Chem. J. **47**, 368–71 (1912).
(83) FRANKLIN, E. C.: J. Am. Chem. Soc. **37**, 847–52 (1915).
(84) FRANKLIN, E. C.: J. Am. Chem. Soc. **37**, 2279–95 (1915).
(85) FRANKLIN, E. C.: J. Am. Chem. Soc. **44**, 492–3 (1922).
(85a) FRANKLIN, E. C.: J. Phys. Chem. **27**, 177 (1923).
(86) FRANKLIN, E. C.: J. Am. Chem. Soc. **55**, 4912–15 (1933).
(87) FRANKLIN, E. C.: J. Am. Chem. Soc. **56**, 568–9 (1934).
(88) FRANKIN, E. C.: *The Nitrogen System of Compounds*, p. 39. Reinhold Publishing Corporation, New York (1935).
(89) FRANKLIN, E. C., AND CADY, H. P.: J. Am. Chem. Soc. **26**, 512 (1904).
(90) FRANKLIN, E. C., AND HINE, T. B.: J. Am. Chem. Soc. **34**, 1497–1501 (1912).

(91) Reference 90, p. 1497.
(92) Franklin, E. C., and Stafford, O. F.: Am. Chem. J. **28,** 83–107 (1902).
Freitag, C.: See Frank, H. H.
(93) Friedricks, F.: J. Am. Chem. Soc. **35,** 244–7 (1913).
Gaddy, V. L.: See Krase, H. J.
Galter, E.: See Philippi, E.
Geisel, E.: See Ruff, O.
(94) Gesellschaft für Kohlentechnik: British patent 436,692; Chem. Abstracts **30,** 2201–2 (1936).
(95) Gesellschaft für Kohlentechnick m. b. H. (Brodkorb, F., Glund, W., and Klempt, W., inventors): German patent 655,563; Chem. Abstracts **32,** 3767 (1938).
(96) Gierut, J. A., Sowa, J. F., and Nieuwland, J. A.: J. Am. Chem. Soc. **58,** 898 (1936).
Gilmore, A. E.: See Bergstrom, F. W.
(97) Glattfeld, J. W. E., and Macmillan, D.: J. Am. Chem. Soc. **56,** 2481–2 (1934).
(98) Glattfeld, J. W. E., and Macmillan, D.: J. Am. Chem. Soc. **58,** 898–901 (1936).
Glund, W.: See Gesellschaft für Kohlentechnik.
(99) Govaert, F.: Natuurw. Tijdschr. **15,** 149–53 (1933); *cf.* Chem. Abstracts **28,** 740 (1934).
(100) Govaert, F. J.: Proc. Acad. Sci. Amsterdam **37,** 156–62 (1934); *cf.* Chem. Abstracts **28,** 4038 (1934).
(101) Grandadam: Compt. rend. **180,** 1598–9 (1925).
Greer, W. N.: See Kraus, C. A.
(102) Groggins, P. H.: Ind. Eng. Chem. **25,** 274–9 (1933).
(103) Groggins, P. H.: *Unit Processes in Organic Syntheses*, pp. 272–323. McGraw-Hill Book Co., New York (1935); 2nd edition, pp. 299–361 (1938).
(104) Groggins, P. H., and Newton, H. P.: Ind. Eng. Chem. **25,** 1030–3 (1933),
(105) Groggins, P. H., and Stirton, A. J.: Ind. Eng. Chem. **25,** 42–9, 169–75 (1933).
(106) Groggins, P. H., and Stirton, A. J.: Ind. Eng. Chem. **29,** 1353–61 (1937).
(107) Reference 106, pp. 1353–4.
(108) Reference 106, pp. 1359–60.
Grove, C.: See Brönsted, J. N.
Gruntfest, I. J.: See Foster, L. S.
Gur'yanova, E. N.: See Monoszon, A. M.
Guttmann, R.: See Voss, W.
(109) Hall, N. F.: Chem. Rev. **8,** 193–4, 198–9, 202 (1931).
Hamblet, C.: See Moureu, H.
(110) Hammett, L. P.: J. Am. Chem. Soc. **50,** 2667 (1928).
(111) Harned, H. S., and Samaras, N. N. T.: J. Am. Chem. Soc. **54,** 1–8 (1932).
(112) Hauser, C. R., and Jordan, E.: J. Am. Chem. Soc. **57,** 2450–6 (1935).
(113) Hauser, C. R., Jordan, E., and O'Connor, R.: J. Am. Chem. Soc. **57,** 2456–8 (1935).
Hawes, W. W.: See Kraus, C. A.
Heierman, F.: See Wibaut, J. P.
(114) Hein, F., and Markert, E.: Ber. **61B,** 2255–67 (1928).
(115) Heisig, G. B.: J. Am. Chem. Soc. **53,** 3245 (1931).
(116) Heisig, G. B., and Davis, H. M.: J. Am. Chem. Soc. **57,** 339 (1935).
(116a) Helferich, B., and Mittag, R.: **71B,** 1588 (1938).

HENDGEN, F.: See Philippi, E.
HENNION, G. F.: See Killian, D. B.
HERNLER, F.: See Philippi, E.
HETHERINGTON, H. C.: See DeRopp, H. W.
HILBERT, G. E.: See Pinck, L. A.
HINE, T. B.: See Franklin, E. C.
(117) HOFMANN, A. W.: Ber. **4**, 265 (1871).
(118) HOWARD, D. H., JR., AND BROWNE, A. W.: J. Am. Chem. Soc. **56**, 2348–57 (1934).
(119) HUGHES, E. D., INGOLD, C. K., AND SCOTT, A. D.: J. Chem. Soc. **1937**, 1201–8.
(120) HUGOT, C.: Compt. rend. **126**, 1719–22 (1898).
(121) HUGOT, C.: Compt. rend. **139**, 54–6 (1904).
(122) HUGOT, C.: Compt. rend. **141**, 1235–7 (1905).
HUGOT, C.: See Vigoroux, E.
HUND, W. J.: See N. V. de Bataafsche Petroleum Maatschappij.
HURD, C. D.: See Meinert, R. N.
(123) HURD, C. D., AND MEINERT, R. N.: J. Am. Chem. Soc. **53**, 291, 296 (1930).
(124) I. G. Farbenindustrie A.-G.: British patent 404,744; Chem. Abstracts **28**, 4068 (1934).
(125) I. G. Farbenindustrie A.-G.: British patent 435,863; Chem. Abstracts **30**, 1385 (1936).
(126) I. G. Farbenindustrie A.-G.: British patent 437,873; Chem. Abstracts **30**, 2208 (1936).
(127) INGOLD, C. K.: J. Chem. Soc. **1930**, 1033–9.
INGOLD, C. K.: See also Hughes, E. D.
(128) INGOLD, C. K., AND MARSHALL, P. G.: J. Chem. Soc. **1926**, 3084.
(129) IRVINE, J. C., AND ROUTLEDGE, D.: J. Am. Chem. Soc. **57**, 1411–14 (1935).
(130) JACKSON, D. R.: Dissertation, The Ohio State University, 1936.
JACOB, A.: See Braun, J. von.
(131) JANSEN, H. E., AND WIBAUT, J. P.: Rec. trav. chim. **56**, 709–13 (1937).
JEANMAIRE, A.: See Schwarz, R.
(132) JELINEK, V. C., AND UPSON, F. W.: J. Am. Chem. Soc. **60**, 356 (1938).
(133) JOANNIS, A.: Compt. rend. **119**, 557–9 (1894).
(134) JOANNIS, A.: Compt. rend. **135**, 1106–9 (1902).
(135) JOANNIS, A.: Compt. rend. **139**, 364–6 (1904).
(136) JOANNIS, A.: Ann. chim. phys. [8] **7**, 101–7 (1906).
(137) JOHNSON, W. C., MOREY, G. H., AND KOTT, A. E.: J. Am. Chem. Soc. **54**, 4278–84 (1932).
(138) JOHNSON, W. C., AND PARSONS, J. B.: J. Phys. Chem. **36**, 2588–91 (1932).
(139) JOHNSON, W. C., AND PECHUKAS, A.: J. Am. Chem. Soc. **59**, 2068–71 (1937).
(140) JOHNSON, W. C., AND SIDWELL, A. E.: J. Am. Chem. Soc. **55**, 1884–9 (1933).
(141) JOHNSON, W. C., AND WHEATLEY, A. C.: Z. anorg. allgem. Chem. **216**, 278 (1934).
(142) JONES, P. C.: U. S. patent 2,046,876; Chem. Abstracts **30**, 5592 (1936).
JORDAN, E.: See Hauser, C. R.
(143) JUZA, R., FASOLD, K., AND KUHN, W.: Z. anorg. allgem. Chem. **234**, 86–8 (1937).
KAHLER, W. H.: See Kraus, C. A.
(144) KARRER, P., AND KLARER, W.: Helv. Chim. Acta **8**, 393 (1925).
KAWAMURA, T.: See Kraus, C. A.
KEYSER, L. S.: See Shriner, R. L.

(145) Killian, D. B., Hennion, G. F., and Nieuwland, J. A.: J. Am. Chem. Soc. **56,** 1384–5 (1934).
(146) Kilpatrick, M., Jr., and Chase, E. F.: J. Am. Chem. Soc. **53,** 1739–44 (1931).
Klarer, W.: See Karrer, P.
Kleinberg, J.: See Audrieth, L. F.
(147) Klemm, W., and Rockstrob, J.: Z. anorg. allgem. Chem. **176,** 181–99 (1928).
Klempt, W.: See Gesellschaft für Kohlentechnik.
Kott, A. E.: See Johnson, W. C.
(148) Kranzfelder, A. L., and Sowa, F. J.: J. Am. Chem. Soc. **59,** 1490–2 (1937).
(149) Krase, H. J., and Gaddy, V. L.: J. Am. Chem. Soc. **52,** 3088–93 (1930).
(150) Krase, H. J., Gaddy, V. L., and Clark, K. G.: Ind. Eng. Chem. **22,** 289–93 (1930).
(151) Kraus, C. A.: J. Am. Chem. Soc. **35,** 1735–7 (1913).
(152) Kraus, C. A.: J. Am. Chem. Soc. **46,** 2196–204 (1924).
(153) Kraus, C. A., and Bien, P. B.: J. Am. Chem. Soc. **55,** 3610 (1933).
(154) Kraus, C. A., and Brown, C. L.: J. Am. Chem. Soc. **52,** 3690–6 (1930).
(155) Kraus, C. A., and Brown, C. L.: J. Am. Chem. Soc. **52,** 4031–5 (1930).
(156) Reference 155, pp. 4031–2.
(157) Kraus, C. A., and Brown, E. H.: J. Am. Chem. Soc. **51,** 2690–6 (1929).
(158) Kraus, C. A., and Brown, E. H.: J. Am. Chem. Soc. **52,** 4414–18 (1930).
(159) Kraus, C. A., and Carney, E. S.: J. Am. Chem. Soc. **56,** 765–8 (1934).
(160) Kraus, C. A., and Eatough, H.: J. Am. Chem. Soc. **55,** 5008 (1933).
(161) Reference 160, pp. 5009–11.
(162) Kraus, C. A., and Flood, E. A.: J. Am. Chem. Soc. **54,** 1638 (1932).
(163) Reference 162, p. 1639.
(164) Reference 162, pp. 1639–41.
(165) Reference 162, pp. 1640–4.
(166) Kraus, C. A., and Foster, L. S.: J. Am. Chem. Soc. **49,** 457–67 (1927).
(167) Reference 166, p. 462.
(168) Kraus, C. A., and Greer, W. N.: J. Am. Chem. Soc. **44,** 2629–33 (1922).
(169) Kraus, C. A., and Greer, W. N.: J. Am. Chem. Soc. **45,** 3078–83 (1923).
(170) Kraus, C. A., and Greer, W. N.: J. Am. Chem. Soc. **47,** 2568–75 (1925).
(171) Reference 170, p. 2569.
(172) Kraus, C. A., and Hawes, W. W.: J. Am. Chem. Soc. **55,** 2776–85 (1933).
(173) Kraus, C. A., and Kahler, W. H.: J. Am. Chem. Soc. **55,** 3537–42 (1933).
(173a) Kraus, C. A., and Kawamura, T.: J. Am. Chem. Soc. **45,** 2756–63 (1923).
(174) Kraus, C. A., and Kurtz, H. F.: J. Am. Chem. Soc. **47,** 57–9 (1925).
(175) Kraus, C. A., and Neal, A. M.: J. Am. Chem. Soc. **51,** 2403–7 (1929).
(176) Kraus, C. A., and Nelson, W. K.: J. Am. Chem. Soc. **56,** 195–202 (1934).
(177) Kraus, C. A., and Rosen, R.: J. Am. Chem. Soc. **47,** 2747 (1925).
(178) Kraus, C. A., and Sessions, W. V.: J. Am. Chem. Soc. **47,** 2361–8 (1925).
(179) Reference 178, p. 2362.
(180) Kraus, C. A., and Sherman, C. S.: J. Am. Chem. Soc. **55,** 4696 (1933).
(181) Kraus, C. A., and Toonder, F. E.: J. Am. Chem. Soc. **55,** 3547–54 (1933).
(182) Kraus, C. A., and Wooster, C. B.: J. Am. Chem. Soc. **52,** 372–6 (1930).
Kuffner, F.: See Späth, E.
Kuhn, W.: See Juza, R.
Kurtz, H. F.: See Kraus, C. A.
(182a) Lake, W. H. G., and Peat, S.: J. Chem. Soc. **1938,** 1417–21.
Lapworth, A.: See Fitzgerald, E.
(183) Lawrence, A. E.: U. S. patent 2,087,980; Chem. Abstracts **31,** 6678 (1937).

(184) LENGFELD, F.: Am. Chem. J. **21**, 531–7 (1899).
(185) LEVENE, P. A., AND MUSKAT, I. E.: J. Biol. Chem. **105**, 435–6 (1934).
LINTNER, J.: See Späth, E.
LOWSON, W.: See Dawson, H. M.
MACMILLAN, D.: See Glattfeld, J. W. E.
(185a) MACMULLIN, R. B. (to The Mathieson Alkali Works): British patent **437,453**; Chem. Abstracts **30**, 2206 (1936).
MACMULLIN, R. B.: See also Cunningham, G. L.
MAHR, C.: See Strecker, W. von.
MAJOR, R. T.: See Addinall, C. R.
MARKERT, E.: See Hein, F.
MARSHALL, P. G.: See Ingold, C. K.
(185b) The Mathieson Alkali Works: French patent 751,699; Chem. Abstracts **28**, 863 (1934).
The Mathieson Alkali Works: See also Cunningham, G. L., and MacMullin, R. B.
(186) MCCHESNEY, E. W., AND MILLER, C. O.: J. Am. Chem. Soc. **53**, 3888–96 (1931).
MCCREADY, W. K.: See Cunningham, G. L.
MCFARLAND, D. F.: See Bartow, E.
MEDVEDEV, S.: See Abkin, A.
MEINERT, R. N.: See Hurd, C. D.
(187) MEINERT, R. N., AND HURD, C. D.: J. Am. Chem. Soc. **52**, 4544 (1930).
(188) MENGE, G. A.: J. Am. Chem. Soc. **56**, 2197 (1934).
(189) MEYER, F.: Compt. rend. **143**, 280–2 (1906).
(189a) MEYER, J., AND BACKA, R.: Z. anorg. Chem. **135**, 189–91 (1924).
MILLER, C. O.: See McChesney, E. W.
(190) MILLER, H. F., AND BACHMAN, G. B.: J. Am. Chem. Soc. **57**, 2444 (1935).
(191) MILLER, R. F., AND ADAMS, R.: J. Am. Chem. Soc. **59**, 1736–8 (1937).
MITCHELL, N. W.: See Wooster, C. B.
MONOSZON, A. M.: See Pleskov, V. A.
(192) MONOSZON, A. M., GUR'YANOVA, E. N., AND SHATENSHTEIN, A. I.: J. Chem. Ind. (Moscow) **12**, 279–81 (1935); *cf.* Chem. Abstracts **30**, 4903 (1935).
MOREY, G. H.: See Johnson, W. C.
MORRISON, A. B.: See White, G. F.
(193) MORSCH, K.: Monatsh. **61**, 299–307 (1932).
(194) MORSCH, K.: Monatsh. **63**, 224–9 (1933).
(195) Reference 194, p. 229.
(196) MORSCH, K.: Monatsh. **64**, 337 (1934).
(197) MOUREU, H., AND HAMBLET, C.: Compt. rend. **200**, 2184–6 (1935).
(198) MOUREU, H., AND HAMBLET, C.: J. Am. Chem. Soc. **59**, 33–40 (1937).
(199) MOUREU, H., AND ROCQUET, P.: Compt. rend. **197**, 1643–5 (1933).
(200) MOUREU, H., AND ROCQUET, P.: Compt. rend. **200**, 1407–10 (1935).
(201) MOUREU, H., AND ROCQUET, P.: Bull. soc. chim. [5] **3**, 821–41 (1936).
(202) MOUREU, H., AND WETROFF, G.: Compt. rend. **201**, 1381–3 (1935).
(203) MUSKAT, I. E.: J. Am. Chem. Soc. **56**, 693–5 (1934).
(204) Reference 203, p. 693.
(205) MUSKAT, I. E.: J. Am. Chem. Soc. **56**, 2449–51 (1934).
(206) Reference 205, pp. 2450–1.
MUSKAT, I. E.: See also Levene, P. A.
NEAL, A. M.: See Kraus, C. A.
NELSON, W. K.: See Kraus, C. A.
NEWTON, H. P.: See Groggins, P. H.

(207) Nienburg, H.: Ber. **68B,** 2232–4 (1935).
Nieuwland, J. A.: See Gierut, J. A., and Killian, D. B.
(208) N. V. de Bataafsche Petroleum Maatschappij: French patent 790,002; Chem. Abstracts **30,** 2987 (1936).
(209) N. V. de Bataafsche Petroleum Maatschappij (assignees of Hund, W. J., and Rosenstein, L.): British patent 445,148; Chem. Abstracts **30,** 6389 (1936).
O'Connor, R.: See Hauser, C. R.
(210) Oda, R.: Sci. Papers Inst. Phys. Chem. Research (Tokyo) **24,** No. 510, 171–3 (1934); *cf.* Chem. Abstracts **28,** 6582 (1934).
(211) Oda, R., and Wada, S.: J. Soc. Chem. Ind. Japan, Suppl. binding **37,** 295–6 (1934); *cf.* Chem. Abstracts **28,** 6005 (1934).
(212) Ôeda, H.: Bull. Chem. Soc. Japan **11,** 385–9 (1936).
Parsons, J. B.: See Johnson, W. C.
(213) Pascal, P., and Dansette, A.: Compt. rend. **190,** 25–7 (1930).
Pechukas, A.: See Johnson, W. C.
Philippi, E.: See Stosius, K.
(214) Philippi, E., and Galter, E.: Monatsh. **51,** 253–66 (1929).
(215) Philippi, E., Hendgen, F., and Hernler, F.: Monatsh. **69,** 281–2 (1936).
(216) Philippi, E., and Spenner, E.: Monatsh. **36,** 97–111 (1915).
(217) Reference 216, p. 106.
(218) Reference 216, p. 107.
(219) Picon, M.: Compt. rend. **175,** 1213–16 (1922).
(220) Picon, M.: Bull. soc. chim. [4] **35,** 979–80 (1924).
(221) Pinck, L. A., and Hilbert, G. E.: J. Am. Chem. Soc. **56,** 490 (1934).
(222) Pinck, L. A., and Hilbert, G. E.: J. Am. Chem. Soc. **59,** 8–13 (1937).
Pinczesi, I.: See Zechmeister, L.
Pinkernelle, W.: See Braun, J. von.
(223) Pleskov, V. A., and Monoszon, A. M.: Acta Physicochim. U.R.S.S. **1,** 725–8 (1935).
(224) Pogosyan, U. P., and Nikeyan, G. M.: Russian patent 50,533; Chem. Abstracts **31,** 8549 (1937).
Popova, A. N.: See Bolotov, B. A.
Pugh, W.: See Thomas, J. S.
(225) Rengade, E.: Compt. rend. **141,** 196–8 (1905).
(226) Roberts, R. G., and others: Proc. Soc. Exptl. Biol. Med. **31,** 998 (1934).
(227) Roberts, R. G., Tweedy, W. R., and Smullen, G. H.: J. Biol. Chem. **112,** 209–14 (1935).
Robinson, W. B.: See Bullard, R. H.
Rockstrob, W.: See Klemm, J.
Rocquet, P.: See Moureu, H.
Rosen, R.: See Kraus, C. A.
Rosenstein, L.: See N. V. de Bataafsche Petroleum Maatschappij.
Rosset: See Besson.
Routledge, D.: See Irvine, J. C.
Ruehle, A. E.: See Cline, J. K.
(228) Ruff, O., and Eisner, F.: Ber. **41,** 2250 (1908).
(229) Ruff, O., and Geisel, E.: Ber. **39,** 842–3 (1906).
(230) Ruff, O., and Treidel, O.: Ber. **45,** 1364 (1912).
Ryan, J. F.: See Wooster, C. B.
Samaras, N. N. T.: See Harned, H. S.
Schenk, P. W.: See Schwarz, R.
Scherer, P. C.: See Chambers, R. W.

Schlösser, P.: See Silesia, Verein Chemischer Fabriken.
Schöpfer, I.: See Wacek, A. von.
(231) Schumb, W. C., and Sundström, R. F.: J. Am. Chem. Soc. **55,** 596–604 (1933).
(232) Schwarz, R.: Angew. Chem. **48,** 221 (1935).
(233) Schwarz, R., and Giese, H.: Ber. **67B,** 1109–15 (1934).
(234) Schwarz, R., and Jeanmaire, A.: Ber. **65B,** 1443 (1932).
(235) Schwarz, R., and Jeanmaire, A.: Ber. **65B,** 1662 (1932).
(236) Schwarz, R., and Schenk, P. W.: Ber. **63B,** 297–300 (1930).
(237) Schwarz, R., and Sexauer, W.: Ber. **59B,** 333–7 (1926).
(238) Schwarz, R., and Striebich, H.: Z. anorg. allgem. Chem. **223,** 399–408 (1935).
Schwarzkopf, H. E.: See Strecker, W. von.
Scott, A. D.: See Hughes, E. D.
Sessions, W. V.: See Kraus, C. A.
Sexauer, W.: See Schwarz, R.
(239) Shatenshtein, A. I.: J. Phys. Chem. (U.S.S.R.) **4,** 703–5 (1933).
(240) Shatenshtein, A. I.: Z. physik. Chem. **A171,** 286–8 (1934).
(241) Shatenshtein, A. I.: Acta Physicochim. U.R.S.S. **3,** 37–52 (1935).
(242) Shatenshtein, A. I.: J. Phys. Chem. (U.S.S.R.) **8,** 613–18 (1936); *cf.* Chem. Abstracts **31,** 2076 (1937).
(243) Shatenshtein, A. I.: J. Phys. Chem. (U.S.S.R.) **8,** 696–708 (1936); *cf.* Chem. Abstracts **31,** 2075 (1937).
(244) Shatenshtein, A. I.: Acta Physicochim. U.R.S.S. **5,** 841–52 (1936).
(245) Reference 244, pp. 846–52.
(246) Shatenshtein, A. I.: J. Am. Chem. Soc. **59,** 432–5 (1937).
Shatenshtein, A. I.: See also Monoszon, A. M.
Sherman, C. S.: See Kraus, C. A.
(247) Shriner, R. L., and Keyser, L. S.: J. Am. Chem. Soc. **60,** 286–8 (1938).
(248) Sidgwick, N. V.: *The Electronic Theory of Valency,* pp. 156–8. Oxford University Press, London (1927).
Sidwell, A. E.: See Johnson, W. C.
(249) Silesia, Verein Chemischer Fabriken (Schlösser, P., inventor): German patent 630,966; 632,130; Chem. Abstracts **31,** 111, 1826 (1937).
(250) Singleton, H. M., and Edwards, W. R., Jr.: J. Am. Chem. Soc. **60,** 542 (1938).
(251) Slobutsky, C., Audrieth, L. F., and Campbell, R. W.: Proc. Natl. Acad. Sci. U. S. **23,** 611–15 (1937).
Smith, F. B.: See Wooster, C. B.
Smith, G. E. P., Jr.: See Blair, J. S.
(252) Smith, G. E. P., Jr., and Bergstrom, F. W.: J. Am. Chem. Soc. **56,** 2095–6 (1934).
Smullen, G. H.: See Roberts, R. G.
Sokolova, Y. K.: See Bolotov, B. A.
Somieski, K.: See Stock, A.
Southwood, W. W.: See Thomas, J. S.
Sowa, J. F.: See Gierut, J. A., and Kranzfelder, A. I.
(253) Spacu, P.: Z. anorg. allgem. Chem. **230,** 181–6 (1936).
(254) Spacu, P.: Z. anorg. allgem. Chem. **232,** 225–8 (1937).
(255) Späth, E., Kuffner, F., and Lintner, J.: Ber. **69B,** 2057 (1936).
(256) Späth, E., and Lintner, J.: Ber. **69B,** 2727–31 (1936).
Spenner, E.: See Philippi, E.
Stafford, O. F.: See Franklin, E. C.
(257) Stähler, A.: Ber. **38,** 2627–9 (1905).

(258) Stähler, A.: Ber. **47,** 909–12 (1914).
(259) Stähler, A., and Denk, B.: Ber. **38,** 2611 (1905).
Stirton, A. J.: See Groggins, P. H.
(260) Stock, A., and Blix, M.: Ber. **34,** 3039–47 (1901).
(260a) Stock, A., and Somieski, K.: Ber. **54,** 740–58 (1921).
(260b) Stock, A., and Zeidler, F.: Ber. **56,** 986–91 (1923).
(261) Reference 260b, pp. 992–4.
(262) Stosius, K., and Philippi, E.: Monatsh. **45,** 457–70 (1924).
(263) Stosius, K., and Philippi, E.: Monatsh. **45,** 569–72 (1924).
(264) Strain, H. H.: J. Am. Chem. Soc. **50,** 2218–23 (1928).
(265) Strain, H. H.: J. Am. Chem. Soc. **52,** 820–3 (1930).
(266) Strain, H. H.: J. Am. Chem. Soc. **52,** 3383–4 (1930).
(267) Strain, H. H.: J. Am. Chem. Soc. **54,** 1221–8 (1932).
(268) Strecker, W. von, and Mahr, C.: Z. anorg. allgem. Chem. **221,** 199–208 (1934).
(269) Strecker, W. von, and Schwarzkopf, H. E.: Z. anorg. allgem. Chem. **221,** 193–8 (1934).
Striebich, H.: See Schwarz, R.
Sundström, R. F.: See Schumb, W. C.
(270) Thomas, J. S., and Pugh, W.: J. Chem. Soc. **1931,** 62–6; *cf.* ibid. **1926,** 1051–61.
(270a) Thomas, J. S., and Southwood, W. W.: J. Chem. Soc. **1931,** 2083–97.
Toonder, F. E.: See Kraus, C. A.
Toth, G.: See Zechmeister, L.
Triedel, O.: See Ruff, O.
Tweedy, W. R.: See Roberts, R. G.
Upson, F. W.: See Jelinek, V. C.
Vasil'ev, B. B.: See Akhumov, E. I.
(271) Vigoroux, E., and Hugot, C.: Compt. rend. **136,** 1670 (1903).
(272) Voss, W., and Guttmann, R.: Ber. **63B,** 1729 (1930).
(273) Wacek, A. von, and Schöpfer, I.: Oesterr. Chem. Ztg. **40,** 63–4 (1937); *cf.* Chem. Abstracts **31,** 3464 (1937).
Wada, S.: See Oda, R.
Wagtendonk, H. M.: See Wibaut, J. P.
Waterman, R. E.: See Cline, J. K.
(274) Watt, G. W.: Thesis, The Ohio State University, 1933.
Watt, G. W.: See also Fernelius, W. C.
Wertheim, E.: See Dains, F. B.
Wetroff, G.: See Moureu, H.
Wheatley, A. C.: See Johnson, W. C.
(275) White, G. F.: J. Am. Chem. Soc. **45,** 779–84 (1923).
(276) White, G. F., Morrison, A. B., and Anderson, E. G. E.: J. Am. Chem. Soc. **46,** 963 (1924).
Wibaut, J. P.: See Jansen, H. E.
(277) Wibaut, J. P., Heierman, F., and Wagtendonk, H. M.: Rec. trav. chim. **57,** 456–8 (1938).
(278) Williams, R. R.: J. Am. Chem. Soc. **58,** 1063–4 (1936).
Williams, R. R.: See also Cline, J. K.
Wirbelauer, W.: See Blix, M.
(279) Wolthorn, H. J.: Thesis, The Ohio State University, 1933.
(280) Wolthorn, H. J.: Dissertation, The Ohio State University, 1934.
Wooster, C. B.: See Kraus, C. A.

(281) WOOSTER, C. B., AND MITCHELL, N. W.: J. Am. Chem. Soc. **52,** 688 (1930).
(282) WOOSTER, C. B., AND RYAN, J. F.: J. Am. Chem. Soc. **54,** 2419 (1932).
(283) WOOSTER, C. B., AND SMITH, F. B.: J. Am. Chem. Soc. **53,** 179–87 (1931).
WORK, R. W.: See Dennis, L. M.
WYNNE-JONES, W. F. K.: See Brönsted, J. N.
(284) YOUNG, R. C.: J. Am. Chem. Soc. **57,** 1195–6 (1935).
(285) ZECHMEISTER, L., AND TOTH, G.: Naturwissenschaften **23,** 35 (1935).
(286) ZECHMEISTER, L., TOTH, G., AND PINCZESI, I.: Ann. **525,** 14–24 (1936).
ZEIDLER, F.: See Stock, A.

SULFAMIC ACID, SULFAMIDE, AND RELATED AQUO-AMMONOSULFURIC ACIDS[1]

L. F. AUDRIETH, M. SVEDA[2], H. H. SISLER[3], AND M. JOSETTA BUTLER[4]

The William Albert Noyes Laboratory, University of Illinois, Urbana, Illinois

Received August 3, 1939

CONTENTS

[1] Presented at the Symposium on the Chemistry of Liquid Ammonia Solutions, held at Milwaukee, Wisconsin, September, 1938, under the auspices of the Division of Physical and Inorganic Chemistry and the Division of Organic Chemistry of the American Chemical Society.

[2] Present address: Grasselli Experimental Laboratory, Cleveland, Ohio.

[3] Present address: Wright Junior College, Chicago, Illinois.

[4] Present address: St. Xavier's College for Women, Chicago, Illinois.

I. Introduction

Prior to the general acceptance of Franklin's work, no successful attempt to systematize the chemistry of nitrogen had ever been made. Recorded investigations, despite their number, lacked a fundamental and unifying basis. It is not surprising, therefore, that the chemical relationships of a large number of nitrogen–sulfur compounds had remained obscure. The recent announcement that sulfamic acid is now being produced on a commercial scale (43, 66) has aroused and stimulated new interest in this group of compounds. In his notable monograph on *The Nitrogen System of Compounds* (59) Franklin calls attention to the fact that a large number of compounds may be regarded as derivatives of a relatively small number of so-called aquo-ammonosulfuric acids. These are depicted diagrammatically in chart I.

Successive replacement of the hydroxyl groups in sulfuric acid by the isosteric NH_2 group leads to sulfamic acid (1) and sulfamide (2), both of which are formally mixed aquo-ammonosulfuric acids. Imidodisulfonic (3) and nitrilosulfonic (4) acids are to be considered deammonation products of sulfamic acid. Imidodisulfamide (5), as well as sulfimide (6) and trisulfimide (7), are in a like manner related to sulfamide. The arrangement depicted in chart I is reminiscent of the relationships which obtain among the aquo-ammonocarbonic acids, such as carbamic acid, urea, cyanic acid, and cyanuric acid. This similarity is apparent not only from a comparison of the respective compounds as listed in table 1, but becomes increasingly more evident in the discussion of the properties and reactions of the various aquo-ammonosulfuric acids which follows.

It is, of course, a relatively simple matter to set up formal relationships among compounds of this sort, as has been done above, and even to extend this limited list of known aquo-ammonosulfuric acids to many other hypothetical ones, but such a procedure deserves no consideration unless (*1*) it represents a definite contribution to the organization of existing knowledge, and (*2*) it is so designed as to be of value in the further investigation of this field. It is, therefore, the purpose of the authors to present in this paper a concise outline of the chemistry of the compounds listed in table 1. An especial attempt is made in the treatment of the subject to demonstrate (*1*) that the aquo-ammonosulfuric acids are prepared (*a*) either by sol-

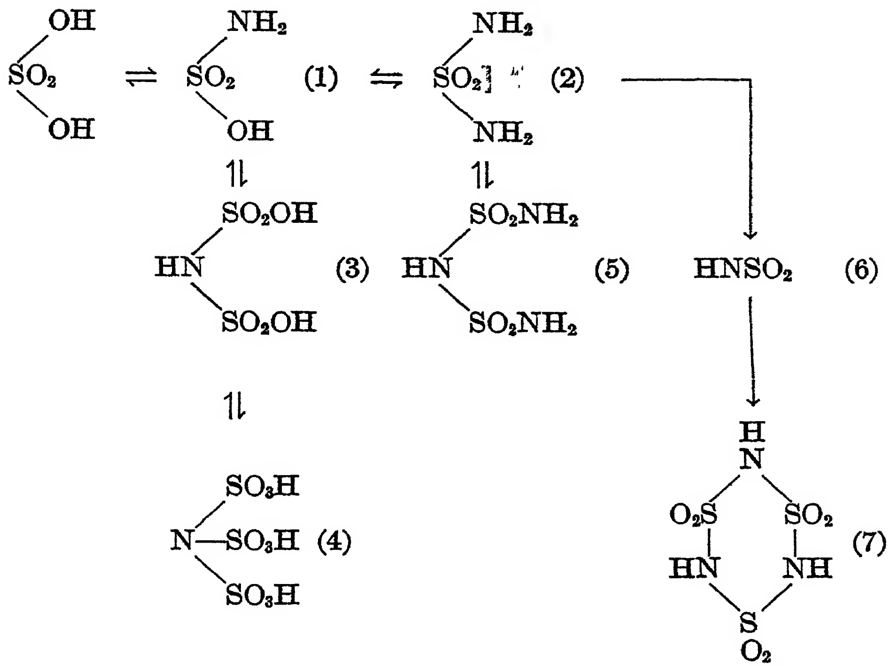


CHART I. The aquo-ammonosulfuric acids. Only known compounds are listed, but it is obvious that the above outline could be extended by processes of ammonation, deammonation and ammonolysis, to include innumerable hypothetical aquo-ammonosulfuric acids and ammonosulfuric acids.

TABLE 1

Mixed aquo-ammono acids related to sulfuric acid and to carbonic acid

RELATED TO $SO_2(OH)_2$		RELATED TO $CO(OH)_2$	
Formula	Name	Formula	Name
H_2NSO_2OH	Sulfamic acid (amidosulfonic, amidosulfuric)	H_2NCOOH	Carbamic acid
$H_2NSO_2NH_2$	Sulfamide (diamide of sulfuric acid, diamidosulfuric acid)	H_2NCONH_2	Urea
$HN(SO_2NH_2)_2$	Imidodisulfamide	$HN(CONH_2)_2$	Biuret
$HN(SO_2OH)_2$	Imidodisulfonic acid	$HN(COOH)_2$	Imidodicarbonic acid
$HNSO_2$	Sulfimide	$HNCO$	Cyanic acid
$(HNSO_2)_3$	Trisulfimide	$(HNCO)_3$	Cyanuric acid

volysis (ammonolysis or aminolysis) of sulfuric acid and its derivatives, or (*b*) in some instances by the nitridation of sulfurous acid; (*2*) that experimental evidence shows that the various aquo-ammonosulfuric acids are actually interconvertible by processes of solvation and desolvation (specifically in the case of the compounds listed in table 1, by the addition or removal of ammonia); (*3*) that hydrolysis leads to sulfuric acid and ammonia (or amines) as end products; and (*4*) that these substances actually do behave as acids in liquid ammonia, warranting their designation as aquo-ammonosulfuric acids.

II. Sulfamic Acid, NH_2SO_3H

A. PREPARATION

Sulfamic acid was first isolated by Berglund (31) in 1878, but it was not until 1887, when Raschig (102) discovered a comparatively simple method for its preparation, that the substance became interesting to chemists. It is referred to as amidosulfonic acid in the older literature, but is now generally known as sulfamic acid because its structure is not unlike that of carbamic acid, NH_2COOH.

Four general methods may be used in the preparation of sulfamic acid. All are based upon certain fundamental reactions which serve to emphasize the chemical character of sulfamic acid as an aquo-ammonosulfuric acid. These are summarized briefly.

1. Ammonolysis of sulfuric acid and related compounds

Such procedures as those which involve the action of aqueous or liquid ammonia upon fluoro- and chloro-sulfonates (132), upon nitrosylsulfuric acid (118), and upon sulfur trioxide addition compounds, *viz.*, with pyridine (9), tertiary amines (11), and dioxane (121), are ammonolytic in character and involve conversion of an aquo derivative of sulfuric acid into an aquo-ammonosulfuric acid.

2. The nitridation of sulfur dioxide, sulfurous acid, sulfites, and hydrosulfites

In some instances sulfamic acid is obtained directly; in other instances related aquo-ammonosulfuric acids are formed which undergo hydrolysis to sulfamic acid. Thus, the Raschig method (102), involving the reaction of sulfur dioxide with hydroxylamine salts in aqueous solution or in the presence of pyridine, may be interpreted as representing the nitridation of tetravalent sulfur to the hexavalent state with the corresponding reduction of the nitrogen to the ammonia state.

$$NH_2OH + SO_2 \rightarrow NH_2SO_2OH$$

Hydroxylamine salts and bisulfites react in a similar fashion to give sulfamates (102, 103, 104).

Acetoxime reacts with sulfurous acid, probably with the intermediate formation of an unstable addition compound, to give sulfamic acid (56a).

$$(CH_3)_2C{=}NOH + SO_2 \rightarrow (CH_3)_2C{=}NSO_3H \xrightarrow{HOH} (CH_3)_2C{=}O + NH_2SO_2OH$$

Chloramine (ammonohypochlorous acid) reacts with sulfites to give sulfamic acid (60).

The interaction of nitrous acid with bisulfite, as depicted in chart II, is a specific case where the reduction of nitrous acid can be shown to proceed through a number of intermediate stages leading, under proper conditions (high temperature and excess bisulfite), to a nitrilosulfonate, a

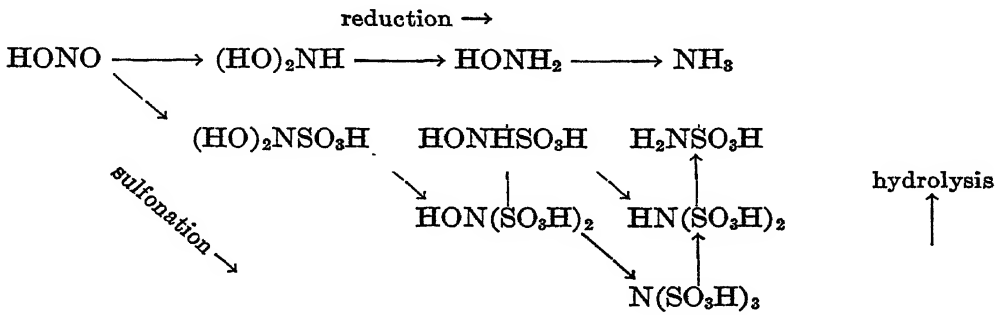


CHART II. Interaction of sulfurous acid with nitrous acid. The reduction of nitrous acid in its simplest aspect is depicted horizontally, leading to ammonia as the end product. Actual, as well as hypothetical, sulfonation reactions are represented diagonally. At low temperature nitrites and bisulfites react to give hydroxylaminodisulfonates, while at higher temperature with excess bisulfite the nitrilosulfonates are obtained. The reactions upward represent hydrolytic processes.

salt of an aquo-ammonosulfuric acid (see page 72). This compound undergoes hydrolysis to the sulfamate.

3. Hydrolysis of aquo-ammonosulfuric acids

Attention has already been called to the fact that the hydrolysis of nitrilosulfonates and imidodisulfonates leads to sulfamic acid (47, 113).

$$N(SO_3K)_3 + 2HOH \rightarrow NH_2SO_2OK + 2KHSO_4$$

$$NH(SO_3K)_2 + HOH \rightarrow NH_2SO_2OK + KHSO_4$$

Hydrolysis of sulfamide in alkaline solution (128) proceeds in accordance with the equation:

$$NH_2SO_2NH_2 + NaOH \rightarrow NH_2SO_2ONa + NH_3$$

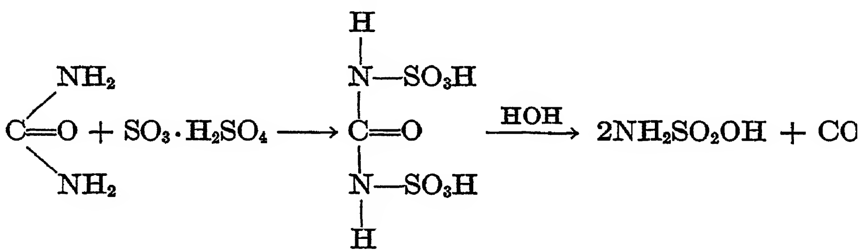
Acid hydrolysis of imidodisulfamide (70) presumably yields both sul acid and sulfamide.

$$HN(SO_2NH_2)_2 + HOH \xrightarrow{H^+} NH_2SO_2OH + NH_2SO_2NH_2$$

All these reactions represent cases of partial hydrolysis of aquo-amr sulfuric acids stopping at the sulfamic acid stage.

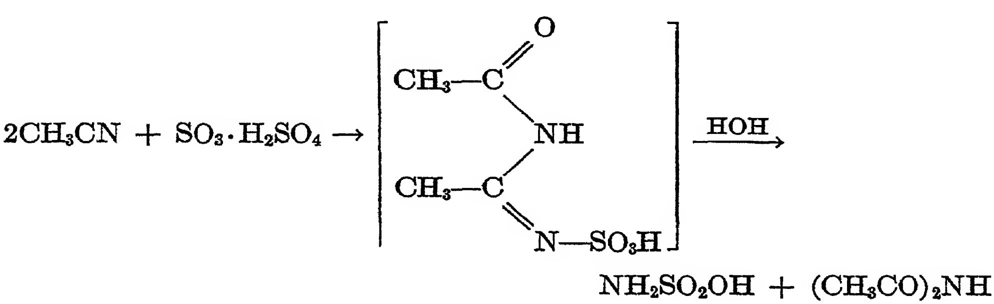
4. Hydrolysis of N-acyl sulfamic acids

Many of these substances are known; some of them are quite resi to hydrolysis. The present commercial method for the producti sulfamic acid, involving the action of fuming sulfuric acid upon ure: 18), presumably goes through the following steps:

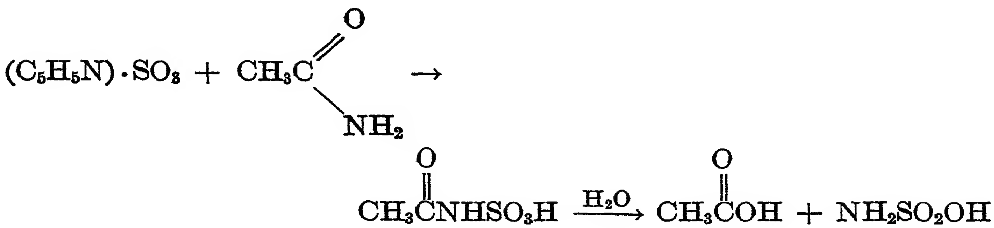


The intermediate is not very stable and under proper conditions unde hydrolysis and decarboxylation.

The action of fuming sulfuric acid on acetonitrile (51) likewise lea an unstable intermediate, which hydrolyzes in accordance with the f ing equation:



Trisubstituted sulfamic acids, such as pyridine–sulfur trioxide, react with acid amides (22, 23) to give *N*-sulfonated products, which hydrolyze to give sulfamic acid.



Chlorosulfonic acid (42, 144) and urea also react to form an unstable *N*-acylsulfamic acid, which likewise breaks down to form sulfamic acid as the principal product.

Cyanic acid (85) and sulfuric acid react in ethereal solution to give sulfamic acid.

B. PHYSICAL PROPERTIES OF SULFAMIC ACID

Sulfamic acid is a crystalline, non-hygroscopic solid, melting with decomposition at 205°C. It is obtained from aqueous solution in the form of rhombic crystals. The solubility in 100 g. of water (43, 66) varies from 14.689 g. at 0°C. to 47.08 g. at 80°C. This solubility in water is decreased markedly by sulfuric acid, so that sulfamic acid is practically insoluble in 70–80 per cent sulfuric acid (43). Its solubility in organic oxygen-containing solvents is only slight or negligible (43). In nitrogenous solvents such as liquid ammonia (37) and formamide (43) sulfamic acid, as might be predicted, is very soluble.

C. CHEMICAL PROPERTIES OF SULFAMIC ACID

Sulfamic acid is highly ionized in aqueous solution. Conductometric (112) and pH measurements (43) place it in the same range of acid strength as hydrochloric, nitric, and sulfuric acids. In view of its strength as an acid, together with its desirable physical properties, sulfamic acid has been recommended for use as a primary standard in acidimetry (38, 71, 72, 92). It has actually been found to be superior to such standards as benzoic acid, succinic acid, potassium biiodate, and potassium acid phthalate. It can be titrated using indicators whose transition points lie within a pH range of 4.5 to 9.

Thus far standard solutions have not been used, because of the fact that sulfamic acid undergoes hydrolysis (that is, conversion from an ammonosulfuric acid into an aquosulfuric acid) in accordance with the equation:

$$NH_2SO_3H + HOH \rightarrow HOSO_3H + NH_3 \text{ (or } NH_4HSO_4\text{)}$$

This reaction is barely discernible at ordinary temperatures (no appreciable concentration of sulfate ion can be detected until after several weeks), but

becomes quite rapid at higher temperatures. According to Cupery (43) a 10 per cent solution of sulfamic acid hydrolyzes to the extent of 40 per cent in 6 hr. at 80°C.

In the cold, chlorine, bromine, and chlorates oxidize sulfamic acid to sulfuric acid. Potassium permanganate, chromic acid, and ferric chloride exert no oxidizing action (47).

At low temperatures hypochlorous acid is said to form an unstable *N*-chloro derivative (135).

$$NH_2SO_2OH + HOCl \rightarrow ClNHSO_3H + HOH$$

The free *N*-chlorosulfamic acid is not sufficiently stable to be isolated. Some of its salts, however, were found to be more stable and were obtained in crystalline form.

Ephraim and Gurewitsch (53) allowed thionyl chloride to react with sulfamic acid at 150°C. in the hope of obtaining sulfamyl chloride. Curiously, other investigators confused the purpose with the results of these experiments, asserting that the work of Ephraim and Gurewitsch could not be substantiated. Actually, these subsequent investigators verified the results obtained by Ephraim and Gurewitsch to the effect that the reaction does not proceed in accordance with the equation (45)

$$NH_2SO_2OH + SOCl_2 \rightarrow NH_2SO_2Cl + HCl + SO_2$$

Denivelle also hoped to prepare sulfamyl chloride by the action of ammonia upon phenyl chlorosulfonate. While it is possible that sulfamyl chloride does form as an intermediate, Denivelle believes that it is decomposed rapidly with the formation of trisulfimide.

$$C_6H_5OSO_2Cl + NH_3 \rightarrow C_6H_5OH + NH_2SO_2Cl$$

$$3NH_2SO_2Cl \rightarrow (NHSO_2)_3 + 3HCl$$

Phosphorus pentachloride was allowed to react with sulfamic acid in another attempt to prepare sulfamyl chloride. A complex addition compound was isolated, corresponding to the formula $PCl_3 \cdot ClSO_2NH_2$ (53). Thus far, all attempts to isolate sulfamyl chloride have failed.

Solutions of sulfamic acid or acid solutions of its salts are rapidly and completely decomposed by the addition of nitrite (32, 66).

$$NH_2SO_2OH + KNO_2 \xrightarrow{H^+} KHSO_4 + N_2 + H_2O$$

This reaction takes place quantitatively and is, therefore, useful in the qualitative and quantitative determination of either sulfamic acid or nitrite, especially in the presence of nitrate (21). This reaction has also been adapted to the separation of lanthanum earths from the rare-earth

metals included in the yttrium group. Sodium nitrite is added to a slightly acid solution of the sulfamates of the rare-earth metals. Reaction takes place, resulting in the precipitation of the double sodium rare-earth sulfates of the lanthanum earths which are practically insoluble, while the corresponding compounds of yttrium and rare earths of higher atomic weight stay in solution (82).

Sulfamic acid reacts with concentrated nitric acid to form nitrous oxide (14).

$$NH_2SO_2OH + HNO_3 \rightarrow H_2SO_4 + H_2O + N_2O$$

This reaction can readily serve as a convenient method for preparing pure nitrous oxide.

D. SALTS OF SULFAMIC ACID (43, 66)

A further proof of the strength of sulfamic acid is furnished by the fact that it reacts readily with basic oxides, hydroxides, and carbonates to yield the corresponding sulfamates. These are, with but few exceptions, soluble in water. The sulfamates of lead, magnesium, and sodium are more soluble in water than the corresponding sulfates, nitrates, chlorides, and acetates. Not only have simple salts of sulfamic acid been isolated, but a number of basic silver and mercury salts as well as a whole series of cobalt (52) and platinum (81, 101) complexes have been characterized.

Ammonium sulfamate is readily obtainable by solution of the free acid in liquid ammonia (37). Solutions of ammonium sulfamate in liquid ammonia are definitely acidic in character. Because of the highly basic nature of ammonia as a solvent, not only is the hydroxylic hydrogen of the acid dissociated, but an amidic hydrogen as well, with the result that sulfamic acid behaves as a dibasic acid forming a disodium (37, 66) and a dipotassium (34) salt.

As might be expected, sulfamic acid enters into combination with compounds containing a basic amino group to give both alkyl- (37, 66) and aryl-substituted ammonium salts.

III. NITROGEN-SUBSTITUTED DERIVATIVES OF SULFAMIC ACID

The large number of *N*-substituted sulfamic acids may readily and easily be placed in certain distinct classes depending upon the nature of the substituent attached to the nitrogen atom. The following classification is recommended. Typical examples are given under each group for illustration.

I. *N*-monosubstituted sulfamic acids, $RNHSO_3H$, where R = alkyl, aryl, acyl, or arylsulfonyl:
   (a) CH_3NHSO_3H (131); $C_{15}H_{31}NHSO_3Na$ (67)
   (b) $C_6H_5NHSO_3K$ (17); α-$C_{10}H_7NHSO_3NH_4$ (99)

(c) $CH_3CONHSO_3K$ (23); $C_{15}H_{31}CONHSO_3Na$ (68)
(d) $C_6H_5SO_2NHSO_3K$ (23)

II. *N*-disubstituted sulfamic acids, $RR'NSO_3H$, where R(R′) = alkyl, aryl, or acyl:

(a) $\left.\begin{array}{l}\text{iso-}C_4H_9\\C_2H_5\end{array}\right\rangle NSO_3H$ (86); $\left.\begin{array}{l}C_{18}H_{37}\\HOC_2H_4\end{array}\right\rangle NSO_3Na$ (67)

(b) $(C_6H_5)_2NSO_3Na$ (76)

(c) $\left.\begin{array}{l}C_6H_5\\C_2H_5\end{array}\right\rangle NSO_3Na$ (130); $\left.\begin{array}{l}p\text{-}CH_3C_6H_4\\CH_3\end{array}\right\rangle NSO_3Na$ (139)

(d) $\left.\begin{array}{l}C_{15}H_{31}CO\\CH_3\end{array}\right\rangle NSO_3Na$ (68)

III. *N*-trisubstituted sulfamic acids (tertiary amine–sulfur trioxide addition compounds), $R(R')(R'')\overset{+}{N}SO_3^-$ (betaine form of $R'R''NSO_3R$):

$(C_2H_5)_3N \cdot SO_3$ (28); $C_5H_5N \cdot SO_3$ (8, 23)

$\left.\begin{array}{l}(CH_3)_2\\C_6H_5\end{array}\right\rangle N \cdot SO_3$ (141)

IV. *N*-alkylidenesulfamic acids, $RCH{=\!=}NSO_3H$ and $R_2C{=\!=}NSO_3H$:

$C_6H_5CH{=\!=}NSO_3H$ (83); $CH_3CH{=\!=}NSO_3K$ (117)
$(CH_3)_2C{=\!=}NSO_3H$ (83)

V. *N*-sulfonated amino acids:

$CH_3CH\left\langle\begin{array}{l}NHSO_3K\\COOK\end{array}\right.$ (24); $\begin{array}{l}CH_2COOK\\|\\NHSO_3K\end{array}$ (24)

Such compounds as CH_3ONHSO_3K (133) might be included as *N*-alkoxysulfamic acids, although they are more properly to be classed as hydroxylamine derivatives. In like manner the compound C_6H_5NHNH-SO_3K (56a) might be classed as an *N*-amino derivative, although it is actually an *N*-sulfonated hydrazine.

A. PREPARATION

The general methods for preparing organic derivatives of sulfamic acid lend themselves to two types of classification. The classification which will be employed here in the discussion of the general methods is based on the type of chemical reaction which occurs during the process. It should be pointed out that there is no essential difference between the methods used in preparing sulfamic acid and those employed in the preparation of the *N*-substituted compounds. Where ammonolytic processes were used, we may now substitute an amine to effect "aminolysis."

A second classification is based on the type of organic compounds used as the starting material and would be employed if it were desired to list and index all known organic derivatives of sulfamic acid. Such a tabula-

tion has been made (124a), but because of its scope cannot be included in this paper.

1. *Aminolytic reactions*

The most useful types of general reactions for the preparation of *N*-substituted sulfamic acids involve the utilization of some form of hexava-

TABLE 2

Aminolysis of some derivatives of aquosulfuric acid

NO.	REACTION	REFERENCE
1	$ArNH_2 + K_2S_2O_7 \rightarrow ArNHSO_3K$	(12)
2	$R_3N + K_2S_2O_7 \rightarrow R_3\overset{+}{N}SO_3^-$	(12)
3	$RR'NH + SO_3 \rightarrow RR'NSO_3H$	(28)
4	$RNH_2 + FSO_3K(Na) \rightarrow RNHSO_3K(Na)$	(67)
5	$RNH_2 + NH_4SO_3F \rightarrow RNHSO_3H$	(131)
6	$ArNH_2 + C_2H_5SO_3Cl \rightarrow Ar\underset{\substack{\vert \\ C_2H_5}}{N}SO_3H$	(130)
7	$R_2NH + R'SO_3Cl \rightarrow R_2NSO_3R'$	(44)
8	$RR'NH + ClSO_3H \rightarrow RR'NSO_3H$	(95)
9	$R_3N + ClSO_3H \rightarrow R_3\overset{+}{N}SO_3^-$	(141)
10	$R_3N + SO_3 \rightarrow R_3\overset{+}{N}SO_3^-$	(86)
11	$R_3N + C_2H_5SO_3Cl \rightarrow R_3\overset{+}{N}SO_3^-$	(8)
12	$RR'NH + ClSO_3H \xrightarrow{R''_3N} RR'NSO_3H$	(105)
13	$ArNH_2$ + dioxane (1 or 2)$SO_3 \rightarrow ArNHSO_3^-$	(121)
14	$R_2NH + NaSO_3Cl \cdot SO_3 \rightarrow R_2NSO_3Na$	(129)
15	$Ar_3N + SO_2Cl_2 \rightarrow$ addition compound $\xrightarrow{ROH} Ar_3\overset{+}{N}SO_3^-$	(10)
16	$R_2NH + SO_3 \xrightarrow{C_6H_5N(CH_3)_2} R_2NSO_3H$	(76)
17	$RSO_4H + ArNH_2 \rightarrow ArNHSO_3H$	(100)
18	$ArCONH_2 + H_2SO_4 \xrightarrow{(CH_3CO)_2O} ArCONHSO_3H$	(36)
19	$RCONH_2 + ClSO_3H \xrightarrow{C_6H_5N(CH_3)_2} RCONHSO_3H$	(68)

lent sulfur for either of the following two fundamental processes: (*a*) the interposition of a bivalent —SO_3— group between the nitrogen and a hydrogen in primary amines, secondary amines, or amides of the types $RCONH_2$ and RSO_2NH_2; and (*b*) the addition of sulfur trioxide through a coördinate link to the nitrogen in a tertiary amine.

It is evident, therefore, that either free sulfur trioxide or a derivative which would readily donate this molecule would be logical sources of the

hexavalent sulfur required for these processes. Indeed, free sulfur trioxide can be used. However, two types of derivatives are usually employed, *viz.*, aquosulfuric acids and ammonosulfuric acids. The use of sulfur trioxide and aquosulfuric acid derivatives is shown in the list of equations in table 2, while the equations involving the use of ammonosulfuric acids are listed in table 3. All of the reactions under these two classifications are aminolytic in character, that is, either sulfur trioxide or one of its derivatives is treated with an amine or an amide. In the equations listed in table 2 both types of fundamental processes mentioned above

TABLE 3

*Aminolysis of some aquo-ammonosulfuric acids**

NO.	REACTION	REFERENCE
20	$\mathrm{RNH_2 + NH_2SO_3H \rightarrow RNHSO_3H \cdot NH_2R}$	(96)
21	$\mathrm{ArNH_2 + NH_2SO_3H \rightarrow ArNHSO_3NH_4}$	(94)
22	$\mathrm{ArNH_2 + C_6H_5\overset{+}{\underset{\underset{(CH_3)_2}{\Vert}}{N}}SO_3^- \rightarrow ArNHSO_3H}$	(141)
23	$\mathrm{ArNH_2 + (CH_3)_3\overset{+}{N}SO_3^- \rightarrow ArNHSO_3H}$	(9)
24	$\mathrm{R_2NH + PySO_3 \rightarrow R_2NSO_3H}$	(138)
25	$\mathrm{ArCONH_2 + NH_2SO_3H \xrightarrow{pyridine} ArCONHSO_3H}$	(88)
26	$\mathrm{RCONH_2 + PySO_3 \rightarrow RCONHSO_3H}$	(23)
27	$\mathrm{ArSO_2NH_2 + PySO_3 \rightarrow ArSO_2NHSO_3H}$	(23)
28	$\mathrm{ArCONHSO_3H \cdot NH_2Ar \xrightarrow{excess\ ArNH_2} ArNHSO_3NH_4}$	(36)
29	$\mathrm{R\underset{\underset{NH_2}{\vert}}{CH}COOH + PySO_3 + K_2CO_3 \xrightarrow{H_2O} R\underset{\underset{NHSO_3K}{\vert}}{CH}COOK}$	(15, 24)

* All of the aquo-ammonosulfuric acids here are sulfamic acids.

are in evidence, while in table 3 the methods of preparation depend solely on the fundamental process designated above under (*a*).

Closer scrutiny of the ammonosulfuric acids used as reactants in the equations in table 3 shows that they are all sulfamic acids themselves. With the exception of the carboxyacylsulfamic acid employed in method 28, these reacting sulfamic acids are of two types: *viz.*, free sulfamic acid itself and trisubstituted sulfamic acids of the general formula $RR'R''NSO_3$. To show the interrelationship of the preparative methods listed in these two tables, all of the sulfamic acids noted as reactants in table 3 are obtained as products of particular reactions in table 2, that is, some sulfamic acids formed by the aminolysis of aquosulfuric acids can themselves be

further aminolyzed to yield other sulfamic acids. Even free sulfamic acid is, of course, obtainable directly or indirectly by the ammonolysis of an aquosulfuric acid derivative.

Occasionally it is difficult to decide whether the reaction occurring involves the aminolysis of a derivative of aquosulfuric acid or the aminolysis of a derivative of ammonosulfuric acid. One of these instances is illustrated in the following equations:

From table 2: $R_3N + ClSO_3H \rightarrow R_3NSO_3$ (equation 9)

$RR'NH + ClSO_3H \rightarrow RR'NSO_3H$ (equation 10)

From table 3: $R_2NH + PySO_3 \rightarrow R_2NSO_3H$ (equation 24)

In reaction 9 a tertiary amine, such as pyridine, is treated with chlorosulfonic acid to give a trisubstituted sulfamic acid, while in reaction 24 a secondary amine is treated with this trisubstituted sulfamic acid to form an *N*,*N*-disubstituted sulfamic acid. In reaction 10 a secondary amine is treated with a mixture of chlorosulfonic acid and a tertiary amine, such as pyridine, and an *N*,*N*-disubstituted sulfamic acid is also the product. Obviously, then, in reaction 10 it is difficult to decide whether the secondary amine reacts with the chlorosulfonic acid, a derivative of aquosulfuric acid, or with the pyridine–sulfur trioxide addition compound, which may form as an intermediate.

The question is clarified somewhat by determining which types of amines yield mono- or di-substituted sulfamic acids by direct aminolysis of chlorosulfonic acid, and which fail to undergo this reaction. An elaboration of equation 8

$$3RR'NH + ClSO_3H \rightarrow RR'NSO_3H \cdot HNRR' + RR'NH \cdot HCl$$

shows that the amine used must be sufficiently basic to form not only the amine salt of the desired sulfamic acid, but also the hydrochloride which is formed as a by-product of the reaction. In general, amines possessing basic properties lower than those of aniline will not yield sulfamic acids when treated directly with chlorosulfonic acid. This failure of certain amines to react with chlorosulfonic acid undoubtedly lies in the inherent instability of the free sulfamic acids formed from amines of low basicity. Therefore, if the amine is not sufficiently basic to stabilize the sulfamic acid by salt formation, aminolysis is not effected. In fact, since hydrogen chloride is more acidic than any of the organic acids under discussion, it can be said that amines which do not form hydrochlorides readily, will not aminolyze chlorosulfonic acid. Thus diphenylamine neither forms a hydrochloride nor aminolyzes chlorosulfonic acid.

However, if a tertiary amine like pyridine or dimethylaniline is added

to a reaction mixture consisting of diphenylamine and chlorosulfonic acid, subsequent treatment with aqueous sodium hydroxide leads to the isolation of the desired sodium diphenylsulfamate. A plausible explanation of the function of the tertiary amine is that the latter forms a salt of the diphenylsulfamic acid when once aminolysis does occur, whereas diphenylamine is not sufficiently basic to stabilize the acid formed.

Several considerations permit a possible conjecture as to whether a trisubstituted sulfamic acid is first formed in this reaction. When a chloroform solution of pyridine is treated with chlorosulfonic acid, the pyridine–sulfur trioxide addition compound precipitates immediately. On the other hand, according to the patent literature, the composite reaction mixture consisting of diphenylamine, pyridine, and chlorosulfonic acid must be heated for some hours on the steam bath before the desired reaction occurs. From these qualitative statements of the reaction rates involved, it is fairly certain that the reaction which finally yields the disubstituted sulfamate is the aminolysis of the trisubstituted sulfamic acid formed as an intermediate.

There is less certainty as to the mechanism of the reaction when the amine used forms a sulfamate from chlorosulfonic acid with or without the addition of a tertiary amine. Methylaniline is an example of this type of amine. Qualitative reaction rates are of no value here, since the reaction times are so nearly alike. Furthermore, the necessary data are not available.

The reactions in which free sulfamic acid is aminolyzed by primary aliphatic or aromatic amines require comment, because the type of sulfamic acid derivative produced appears to depend on the basicity of the amine. This is evident from an examination of equations 20 and 21 in table 3.

$$RNH_2 + NH_2SO_3H \rightarrow RNHSO_3H \cdot H_2NR \quad \text{(equation 20)}$$

$$ArNH_2 + NH_2SO_3H \rightarrow ArNHSO_3NH_4 \quad \text{(equation 21)}$$

In equation 20, R is an aliphatic group such as isoamyl, whereas in equation 21, Ar refers to a strictly aromatic group such as phenyl. If an amine containing either type of organic residue is treated with free sulfamic acid, the product first formed is the corresponding amine sulfamate. When this salt is heated with an excess of the corresponding amine, an *N*-substituted sulfamic acid is formed, undoubtedly with the elimination of a molecule of ammonia. Here, however, the similarity of the reactions is terminated, and the differentiation appears to be a function of the basicity of the amine. Isoamylamine has a basicity comparable to that of ammonia. Therefore, a severalfold excess of the former leads to the forma-

tion of the amine salt of the *N*-substituted sulfamic acid, rather than to the ammonium isoamylsulfamate. On the other hand, the low basicity of aniline permits the preferential formation of ammonium phenylsulfamate.

2. Nitridation reactions

Although the most useful reactions for the preparation of *N*-substituted sulfamic acids are aminolytic in nature, methods exist in which nitridation of some form of sulfur not in the hexavalent state is involved. From an examination of the equations listed in table 4 it is easily discernible that

TABLE 4

Nitridation of some form of sulfur which is not in hexavalent state

NO.	REACTION	REFERENCE
30	$ArNH_2 + SO_2 \xrightarrow[\text{and heat}]{\text{sealed tube}} ArNHSO_3H$	(80)
31	$ArNO_2 + NaHSO_3 \rightarrow ArNHSO_3Na$	(139)
32	$ArNO_2 + Na_2S_2O_4 \rightarrow ArNHSO_3Na$	(119)
33	$ArNO_2 + (NH_4)_2SO_3 \rightarrow ArNHSO_3H$	(99)
34	$RR'R''N{\rightarrow}O + SO_2 \rightarrow RR'R''\overset{+}{N}SO_3^-$	(86)
35	$RR'C{=}NOH + (NH_4)_2SO_3 \rightarrow RR'CHNHSO_3H$	(1)
36	$ArCH{=}NOH + NaHSO_3 \rightarrow Ar\underset{\displaystyle SO_3Na}{\underset{\vert}{C}}HNHSO_3Na$	(97)
37	$ArCH_2NHOH{\cdot}HCl + SO_2 \rightarrow ArCH_2NHSO_3H$	(115)
38	$(ArCH_2)_2NOH + NaHSO_3 \xrightarrow{\text{then add HCl}} (ArCH_2)_2NSO_3H$	(115)
39	$RR'NOH + SO_2 \rightarrow RR'NSO_3H$	(86)
40	$ArNHOH + ArN{=}SO \rightarrow ArNHSO_3H$	(91)
41	$ArN{=}NAr + NH_4HSO_3 \rightarrow NH_2Ar{-}ArNHSO_3NH_4$	(123)
42	$ArN{=}NAr + NaHSO_3 \rightarrow NH_2Ar{-}ArNHSO_3Na$ (small yield)	(35)

the usual nitridizing agents are aromatic nitro compounds and hydroxylamine derivatives. Mono-, di-, and tri-substituted sulfamic acids are formed when hydroxylamines are employed, while only monosubstituted sulfamic acids result from aryl nitro compounds.

When the reactions by which certain types of sulfamic acids can be prepared from free hydroxylamine and ammonia are summed up, as in chart III, certain similarities and differences appear in the types of chemical reactions involved. For instance, alkylation of either parent substance leads to exactly analogous compounds, if the uniformly higher state of oxidation of the hydroxylamines is borne in mind. This difference in oxidation level is important, for on it depends the subsequent choice of the

types of sulfur compounds which lead to the production of identical types of sulfamic acids from the corresponding substituted hydroxylamines and amines.

To determine logically what types of sulfur compounds, when treated with a member of either series of nitrogen compounds, would produce sulfamic acids, the characteristic functional group, $>NSO_3H$, must also be envisioned. It is obvious that when amines are employed the sulfur compound must furnish all three of the oxygen atoms required to produce

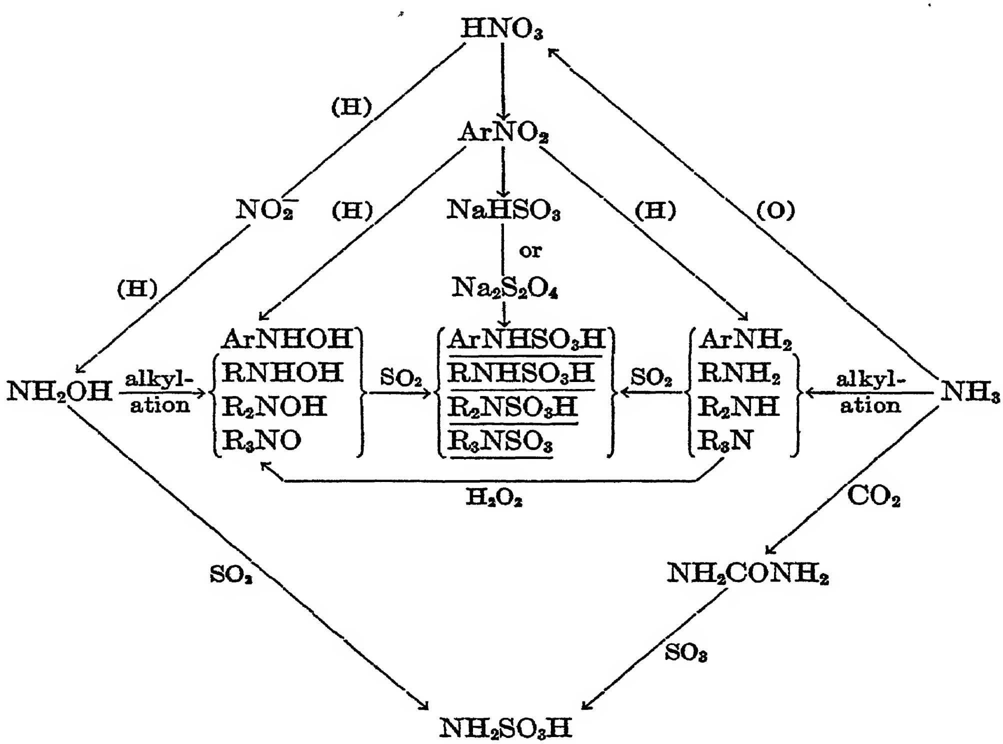


CHART III

this group. On the other hand, hydroxylamines, since they already contain an atom of oxygen, would require treatment with a sulfur compound which need furnish only two atoms of oxygen. Therefore, as has been mentioned under the discussion of the equations in table 2 and table 3, hexavalent sulfur in the form of free sulfur trioxide, or one of its derivatives, is the logical compound of sulfur with which amines can be treated to produce sulfamic acids. Equally logical is the choice of tetravalent sulfur in the form of sulfur dioxide, or one of its derivatives, for the production of sulfamic acids from hydroxylamines. Since sulfamic acids contain

hexavalent sulfur, the overall reaction, by which these compounds are produced from hydroxylamines and a reagent in which sulfur is originally in the tetravalent condition, is one involving nitridation of sulfur.

The interrelation of these methods is evident from the alternative procedures by which the identical sulfamic acids can be prepared from a tertiary amine. The latter can either be treated directly with a derivative of sulfur *trioxide*; or the amine can first be converted to an amine oxide, which can then be treated with sulfur *dioxide*. Both procedures lead to the identical trisubstituted sulfamic acid. Thus, trimethyl- and tripropylsulfamic acids are produced either by the reaction of the corresponding amines with sulfur trioxide, or by the reaction between the amine oxides and sulfur dioxide.

While the production of sulfamic acids from hydroxylamines probably involves only one nitridation process, the formation of sulfamic acids from aromatic nitro compounds probably involves two. When alkali sulfites or hyposulfites are nitridized in aqueous solution by aromatic nitro compounds, the nitrogen in the latter is undoubtedly reduced first to some lower valence state. The resulting aromatic nitrogen compound then reacts with some form of sulfur present in the solution to produce the monosubstituted sulfamic acid. At first glance it is difficult to determine whether the reduction proceeds to the amine stage, with the subsequent aminolysis of the sulfur trioxide derivative formed by the nitridation process, or whether the reduction ceases at the hydroxylamine stage, the sulfamic acid then being formed by another nitridation process involving the excess of the sulfur dioxide derivative present. As is evident from chart III, either process might conceivably occur, since, for example, both aniline and phenylhydroxylamine, when treated with sulfur trioxide and sulfur dioxide, respectively, are converted into phenylsulfamic acid. However, aqueous alkali sulfates, the sulfur trioxide derivatives formed by the first nitridation reaction, do not react with amines, whereas either aqueous sulfur dioxide or an alkali bisulfite does react with phenylhydroxylamine to produce phenylsulfamic acid. Therefore aromatic nitro compounds, when treated with aqueous alkali bisulfites or hyposulfites, are first reduced to the corresponding hydroxylamine, the latter then nitridizing the excess of sulfur dioxide present in order to produce the monoarylsulfamic acid.

3. Hydrolysis of organic derivatives of other aquo-ammonosulfuric acids

The equations listed in table 5 are of interest because Nos. 43, 45, and 46 represent the only known methods for the preparation of polymethylenedisulfamic acids, alkoxysulfamic acids,[5] and polymethylenedioxydisulfamic acids,[5] respectively. Only one example of each of these various types of

[5] These compounds could be considered as derivatives of hydroxylaminomonosulfonic acid.

sulfamic acids is known, the substituent groups being ethylene, methoxy, and ethylenedioxy. It should be noted that the first step in the production of sulfamic acids by the reactions listed in table 5 is the formation of *N*-substituted imidodisulfonic acids. The desired sulfamic acids are then obtained by the removal of one sulfonic acid group through acidic hydrolysis.

4. Miscellaneous methods

Two general methods by which sulfamic acids are produced from ammonosulfuric acids have already been discussed. In the reactions listed in table 3, ammonosulfuric acids which are also sulfamic acids are aminolyzed to produce other sulfamic acids. In those listed in table 5, which have just been discussed, ammonosulfuric acids other than sulfamic acids

TABLE 5

Hydrolysis of other aquo-ammonosulfuric acids

NO.	REACTION	REFERENCE
43	$Br(CH_2)_nBr + KN(SO_3K)_2 \rightarrow (KSO_3)_2N(CH_2)_nN(SO_3K)_2 \xrightarrow{H^+} KSO_3NH(CH_2)_nNHSO_3K$	(136)
44	$R_2SO_4 + KN(SO_3K)_2 \rightarrow RN(SO_3K)_2 \xrightarrow{H^+} RNHSO_3K$	(136)
45	$R_2SO_4 + KON(SO_3K)_2 \rightarrow RON(SO_3K)_2 \xrightarrow{H^+} RONHSO_3K$	(133)
46	$Br(CH_2)_nBr + KON(SO_3K)_2 \rightarrow (KSO_3)_2NO(CH_2)_nON(SO_3K)_2 \xrightarrow{H^+} KSO_3NH{-}O{-}(CH_2)_n{-}O{-}NHSO_3K$	(133)

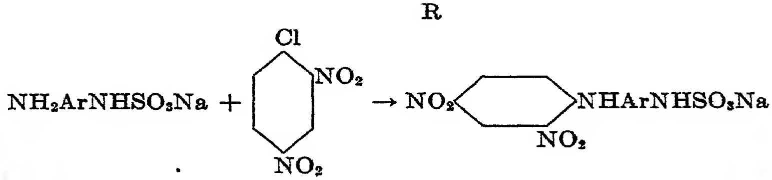
are first alkylated to produce imidodisulfonic acid derivatives, which upon hydrolysis yield the desired sulfamic acids.

In the equations listed in table 6, a number of sulfamic acid derivatives are treated with various reagents in order to produce other derivatives of sulfamic acid. The reactions involved in these transformations are not aminolytic. Furthermore, the ammonosulfuric acids used as starting materials are in themselves sulfamic acids. All other preparative methods are, therefore, grouped together in table 6. Some semblance of classification according to the type of chemical reaction involved is attained by distributing the equations under the five subdivisions.

Several equations in this miscellaneous class are of distinct interest. For instance, equations 48 and 64 illustrate two methods of producing an arylsulfamic acid containing a free amino group in the ring. In addition,

TABLE 6

*Miscellaneous methods for preparation of N-substituted sulfamic acids**

NO.	REACTION	REFERENCE
	a. Hydrolysis	
47	$R_2NSO_2Cl + H_2O \rightarrow R_2NSO_3H$	(26)
48	$RCONHArNHSO_3Na \xrightarrow{NaOH} NH_2ArNHSO_3Na$	(140)
49	$Ar_3\overset{+}{N}SO_3^- \xrightarrow{NaOH}$ decomposition products	(8)
	b. Alcoholysis	
50	$RR'NSO_2Cl + NaOR'' \rightarrow RR'NSO_3R''$	(137)
	c. *N*-acylation	
51	$RCOCl + R'NHSO_3H \xrightarrow{pyridine} RCON\begin{matrix} SO_3H \\ R' \end{matrix}$	(68)
	d. *N*-alkylation or *N*-arylation	
52	$ArCHCl_2 + NH_2SO_3H \rightarrow ArCH{=}NSO_3H$	(83)
53	$ArCH_2Cl + AgSO_3NH_2 \rightarrow ArCH_2NHSO_3CH_2Ar$	(137)
54	$RI + AgSO_3NH_2 \rightarrow R_2NSO_3R$	(137)
55	$CH_3I + \left\{\begin{matrix} NH_2SO_3Ag \\ CH_3NHSO_3Ag \\ (CH_3)_2NSO_3Ag \end{matrix}\right\} \rightarrow (CH_3)_3\overset{+}{N}SO_3^-$	(137)
56	$RCHN_2 + NH_2SO_3H \rightarrow (RCH_2)_3\overset{+}{N}SO_3^-$	(137)
57	Aldehyde patents; reaction is probably as follows: $RCHO + NH_2SO_3(metal) \rightarrow RCH{=}NSO_3(metal)$	(117)
58	$ArNHSO_3Na + RSO_2Cl \xrightarrow{NaOH} Ar\underset{R}{N}SO_3Na$	(130)
59	$ArNHSO_3Na + R_2SO_4 \xrightarrow{NaOH} Ar\underset{R}{N}SO_3Na$	(139)
60		(140)
	e. Reactions involving introduction of substituents on an aryl ring	
61	$(ArNHSO_3)_2Ba + Br_2 \rightarrow$ ring-brominated product	(123)
62	$Ar(NHSO_3H)_2 + HONO \rightarrow XN_2ArNHSO_3H$	(114)
63	$NH_2ArNHSO_3H + HONO \rightarrow XN_2ArNHSO_3H$	(78)
64	$NO_2ArNHSO_3H \xrightarrow{reduction} NH_2ArNHSO_3H$	(78)

* These reactions involve treatment of a sulfamic acid derivative with various agents to effect conversion into other types of organic derivatives.

the first of these two equations focuses attention on the stability of sulfamates towards alkaline solutions, and the second is important in the dye industry. In fact, equations 64, 62, and 63 represent processes which are the basis for several patents in this field. All involve the production of dyes which are prepared by two successive diazotizations. In equation 62 an aromatic diamine which has been converted to the disulfamic acid derivative is treated with sufficient nitrous acid to diazotize only one sulfamic acid group. With one amino group protected as a sulfamate, coupling occurs only on the diazotized portion of the molecule. After this first coupling reaction is complete, the other sulfamic acid group is diazotized by an excess of nitrous acid, and another group is introduced, again by coupling.

B. PHYSICAL PROPERTIES

All metallic and amine salts of *N*-substituted sulfamic acids are solids. The free acids, where they are known, are also solids, including the trisubstituted derivatives which exist in the betaine-like form, $RR'R''NSO_3$. Trisubstituted sulfamic acids in the ester form, $RR'NSO_3R''$, are oils, but on rearrangement to the betaine-like structure, of course, become solids.

C. REACTIONS

1. With alkalies

As a rule, salts of sulfamic acids of the types $RR'NSO_3H$, $ArNHSO_3H$, $RNHSO_3H$, and $ArRNSO_3H$ are fairly stable towards alkaline solutions, whereas compounds of the type $RCH{=}NSO_3M$ (M = metal) are not.

Some monoaryl derivatives like sodium phenylsulfamate and sodium *p*-tolylsulfamate are said to be stable towards fused sodium hydroxide up to 250°C., but decomposition occurs at 280°C. These two sulfamates are also stable towards a tenfold quantity of 50 per cent potassium hydroxide for 1 hr. under reflux. On the other hand, sodium *p*-ethoxyphenylsulfamate undergoes decomposition under the same conditions to the free amine in a short time (139).

This stability towards alkaline solutions is utilized in the preparation of arylsulfamic acids containing a free amino group (140).

$$p\text{-}CH_3CONHC_6H_4NHSO_3Na \xrightarrow{NaOH} p\text{-}NH_2C_6H_4NHSO_3Na$$

N-alkylation, which requires a basic reaction medium, is also possible (139).

$$p\text{-}CH_3C_6H_4NHSO_3Na + (CH_3)_2SO_4 \xrightarrow{NaOH} p\text{-}CH_3C_6H_4\underset{\displaystyle CH_3}{\underset{|}{N}}SO_3Na$$

Trisubstituted sulfamic acids are not stable toward alkaline solutions. Decomposition products have been isolated from the pyridine–sulfur trioxide (16, 8) and the isoquinoline–sulfur trioxide addition compounds (25). In both instances ring fission occurs.

2. With water

As a rule, salts of sulfamic acids which are stable toward alkalies are stable in neutral solution. However, sulfamic acids as a class are more stable toward basic solutions than toward acidic solutions. If hydrolysis proceeds to any extent, hydrogen ion is liberated.

$$RNHSO_3Na + H_2O \rightarrow RNH_2 + NaHSO_4$$

The slight amount of hydrogen ion liberated may catalyze further hydrolysis, and, whereas that slight acidity would be neutralized immediately in alkaline solution, this would not be true in a neutral solution. For this reason, if incipient hydrolysis occurs in a neutral solution, the remaining sulfamate may undergo rapid decomposition.

3. With acids

As mentioned in the preceding section, sulfamic acids as a class are not as stable in acidic solution as they are in basic solution. However, the behavior of alkyl- and aryl-sulfamic acids toward acids is sharply differentiated. No *N*-substituted sulfamic acid in which the nitrogen of the functional group is directly attached to a carbon in an aryl ring is stable for more than a few seconds toward a boiling dilute mineral acid solution. In fact, this instability is employed as a test to differentiate *N*-arylsulfamic acids from the isomeric aminoarylsulfonic acids. A solution containing the suspected *N*-arylsulfamate and some barium chloride is merely acidified with a drop of hydrochloric acid. If on boiling this solution, barium sulfate precipitates, a positive test for an *N*-arylsulfamate is in evidence. The isomeric aminoarylsulfonic acids, of course, do not undergo decomposition under these conditions. Since *N*-arylsulfamic acids are so labile toward acidic solutions, it is a foregone conclusion that very few of them are known in the free state.

On the other hand, both mono- and di-alkylsulfamic acids are much more stable toward acidic solutions, and many of these are known in the free state.

The alkylidenesulfamic acids, that is, compounds of the type $RCH{=}NSO_3$metal, are not stable toward acids (117).

4. With nitrous acid and nitrites

Free acids of the type $RNHSO_3H$, where R is either an alkyl or an aryl group, dissolve in a cold concentrated solution of alkali nitrite. On

standing, salts of the type $RN(NO)SO_3M$ precipitate from solution. These *N*-nitroso derivatives as a class are extremely reactive and sometimes explode when heated (95, 96, 131).

The use of arylsulfamic acids in the dye industry has been discussed under the general methods listed in table 6.

5. Rearrangements

There is a more or less general rule in organic chemistry that almost any functional group on an amino nitrogen directly attached to a carbon in an aryl ring can be rearranged into the aryl ring. The familiar rearrangement of aniline sulfate to sulfanilic acid, and of *N*-nitrosomethylaniline to *p*-nitrosomethylaniline are examples of this rule. The rearrangements of *N*-arylsulfamic acids to the isomeric *p*- and *o*-sulfanilic acids are further examples of the generality of this statement. Thus, potassium phenylsulfamate in a solution containing 20 drops of concentrated sulfuric acid in 25 cc. of ice water yields potassium *o*-sulfanilate after 80 hr., and the latter compound in 25 g. of concentrated sulfuric acid at 180–190°C. yields potassium *p*-sulfanilate after 7 hr. (3). This would lead one to generalize that at a low temperature phenylsulfamic acid rearranges to *o*-sulfanilic acid, and at a high temperature to *p*-sulfanilic acid. The barium salt of phenylsulfamic acid on heating at 180°C. is rearranged to barium *p*-sulfanilate (2, 125). Ammonium salts of *N*-arylsulfamic acids rearrange on heating to form ammonium salts of the isomeric aminoarylsulfonic acids.

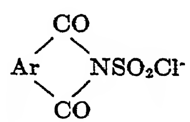
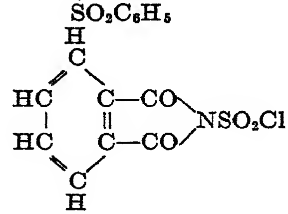
Various workers have postulated that in the Piria reaction, that is, the action of an alkali bisulfite on an aromatic nitro compound, the aminoarylsulfonic acid formed as a by-product is formed from the sulfamic acid by rearrangement. However, some recent work has shown that treatment of the sodium salts of various arylsulfamic acids with hydrochloric acid in concentrations from 0.47 to 5.35 *N* at 47–50°C. permits the quantitative precipitation of the sulfur as barium sulfate. This, of course, precludes any possibility of rearrangement (73). This work is not conclusive, since some investigators carry out the Piria reaction under reflux (139). Obviously, then, checking the rearrangement at 47–50°C. and then stating that in the Piria reaction the aminoarylsulfonic acids are not formed through the *N*-arylsulfamic acids as intermediates, is an error. The only logical conclusion from these studies (73) is that if the Piria reaction is carried out at 47–50°C. rearrangement does not occur.

The esters of *N*-dialkylsulfamic acids are very labile oils. If the three groups are small, as in $(CH_3)_2NSO_3C_2H_5$ and $(CH_3)_2NSO_3CH_3$, the esters rearrange very easily on heating to the betaine-like structure $RR'R''\overset{+}{N}SO_3^-$. Compounds having the latter structure are solids. The com-

pound $(C_2H_5)_2NSO_3C_2H_5$ does not undergo this rearrangement (137, 141). However, the betaine-like form can be prepared from $(C_2H_5)_3N$ and SO_3 (28). Only one ester of an *N*-monosubstituted sulfamic acid, $C_6H_5CH_2$-$NHSO_3CH_2C_6H_5$, is recorded in the literature. This does not rearrange on heating (137).

IV. Nitrogen-substituted Sulfamyl Halides

It is perhaps significant that *N*-substituted sulfamyl halides are quite stable and readily prepared, whereas the parent substance, H_2NSO_2Cl, has not yet been prepared or identified. Four distinct types of sulfamyl chlorides are listed in the literature. The generic name adopted for this survey, the type formula, and a specific example of each follow:

GENERIC NAME	TYPE FORMULA	EXAMPLE	REFERENCE
1. Dialkyl..........	$RR'NSO_2Cl$	$(CH_3)_2NSO_2Cl$	(45)
2. Aryl-acyl	$ArNSO_2Cl$ COR	$C_6H_5NSO_2Cl$ $COCH_3$	(90)
3. (Aryl)-arylsulfonyl...........	$ArNSO_2Cl$ SO_2Ar	$C_6H_5NSO_2Cl$ $SO_2C_6H_5$	(6)
4. Diacyl			(4)

In addition to the various classes of sulfamyl chlorides given above, a number of dialkylsulfamyl fluorides are described in the patent literature (75). Since the information concerning these is very indefinite, little can be said about their properties.

Although the $>NSO_2Cl$ group is common to all four classes of sulfamyl chlorides, the various types of organic residues in the molecule cause a profound difference in many of the reactions of each class toward a particular reagent or set of conditions. Therefore, wherever these peculiarities are in evidence, the reactions of each class will be considered separately.

A. PREPARATION

All sulfamyl chlorides have been prepared by the action of sulfuryl chloride on the appropriate amine or amine hydrochloride, or on a metallic

salt of the appropriate amide. In addition, diethylsulfamyl chloride has been prepared by the action of excess phenyl chlorosulfonate on diethylamine.

1. Preparation of dialkylsulfamyl chlorides

Diethyl- and dimethyl-sulfamyl chlorides are prepared by heating the corresponding amine hydrochloride with an excess of sulfuryl chloride on the steam bath for several hours. After washing the crude product with sodium carbonate solution, the pure compound is obtained by vacuum distillation.

The dialkylsulfamyl chlorides containing larger carbon residues are obtained by treating the free amine with sulfuryl chloride in an indifferent diluent.

An alternative procedure for the preparation of the diethyl derivative depends on the use of excess phenyl chlorosulfonate on the free amine (45).

2. Preparation of sulfamyl chlorides containing either acylsulfonyl or arylsulfonyl residues

The ring-chlorinating action which sulfuryl chloride exerts on free aryl amides is suppressed by utilizing the corresponding alkali salts. Since benzoylphenylsulfamyl chloride is obtained in low yield by this method, a nitrogen-Grignard derivative has been employed with greater success for the preparation of this compound (6).

The procedure requiring either type of metallic derivative presents no technical difficulties, save that of rigorously excluding moisture. The anhydrous alkali salts of the amides are obtained by treating a solution of the amide in toluene either with the finely divided alkali metal or with the alkali ethoxide. For the preparation of the Grignard derivative of benzanilide, the latter must be treated with propylmagnesium chloride, since ethylmagnesium bromide does not give satisfactory results. Either type of metallic derivative is then treated with sulfuryl chloride, and the pure sulfamyl chloride is obtained by recrystallization from an appropriate solvent.

B. PHYSICAL PROPERTIES

The dialkylsulfamyl chlorides are oily liquids, whereas those having at least one acylsulfonyl, or arylsulfonyl, group attached to the nitrogen are solids, regardless of what the other substituent on the nitrogen may be. Without exception, all of the sulfamyl chlorides are practically insoluble in cold water. Boiling water reacts to give decomposition products, presumably hydrolytic in character. Data are available for the solubilities in alcohol of all types of sulfamyl chlorides, with the exception of the

diacyl derivatives. Alcohol is the best crystallizing solvent for the (aryl)-arylsulfonyl derivatives. Aryl-acyl derivatives undergo reaction with the solvent, while the dialkylsulfamyl chlorides probably merely dissolve. Ether and benzene are used as reaction media for the preparation of most sulfamyl halides, and appear, therefore, to be without action.

C. REACTIONS

Since the various classes of sulfamyl chlorides have been studied by investigators for widely differing purposes, they have not all been subjected to the same set of conditions or reagents. Therefore, no mention of a particular class of sulfamyl chlorides will be made in the subdivisions below unless that class has been subjected to the conditions under discussion.

1. Action of heat

Gentle warming has no effect on any of the sulfamyl chlorides. Since the dialkyl derivatives are liquids at room temperature, they are best purified by distillation under reduced pressure. Heating the (aryl)-arylsulfonylsulfamyl chlorides on an oil bath results in the liberation of hydrogen chloride and sulfur dioxide, and the formation of unidentified colored products. When the (aryl)-acylsulfamyl chlorides are heated above their melting points, decomposition also occurs with the formation of hydrogen chloride, sulfur dioxide, and the chloride of the acyl residue in the original molecule. Thus, acetylphenylsulfamyl chloride liberates acetyl chloride as one of the products.

2. Action of hydrolyzing agents

(*a*) *Boiling water.* Dialkylsulfamyl chlorides are rapidly split by boiling water with the formation mainly of the corresponding dialkylsulfamic acid. The hydrochloric acid liberated in this reaction presumably catalyzes to a slight degree the secondary hydrolysis of the sulfamic acid to the corresponding amine and sulfuric acid.

$$RR'NSO_2Cl + H_2O \longrightarrow RR'NSO_3H + HCl$$

$$RR'NSO_3H + H_2O \xrightarrow{HCl} RR'NH \cdot H_2SO_4$$

The first equation above represents one of the general methods of preparing free dialkylsulfamic acids (26).

For a given aryl group in (aryl)-acylsulfamyl chlorides, the reactivity of this class of compounds toward boiling water depends on the acyl group present. Thus, (*1*) phenylformylsulfamyl chloride, (*2*) phenylacetylsulfamyl chloride, and (*3*) phenylbenzoylsulfamyl chloride undergo complete solution in boiling water in 10, 60, and 480 min., respectively.

Titration with standard alkali shows—assuming that hydrochloric acid and sulfuric acid are formed as two of the products—that (*1*) yields formanilide, which is then partially split to aniline and formic acid, (*2*) yields aniline and acetic acid, and (*3*) yields benzanilide (90).

$$\begin{array}{l} \mathrm{C_6H_5NSO_2Cl} \\ \quad\;\, | \\ \quad \mathrm{COR} \end{array} \xrightarrow{\mathrm{H_2O}} \mathrm{HCl} + \mathrm{H_2SO_4} + \begin{array}{l} \mathrm{C_6H_5NH} \\ \quad\;\, | \\ \quad \mathrm{COR} \end{array}$$

(The anilide may undergo further hydrolysis to give aniline and a carboxylic acid.)

(Aryl)-arylsulfonylsulfamyl chlorides are only slowly hydrolyzed by boiling water.

(*b*) *Alkalies.* The reactivity of the (aryl)-acylsulfamyl chlorides toward alkalies again depends on the acyl group present. Thus dilute sodium hydroxide splits phenylformylsulfamyl chloride even in the cold with the formation of benzonitrile, the latter being recognized by its odor (90). Warm alkali must be employed to cause hydrolysis of phenylacetylsulfamyl chloride according to the following equation:

$$\begin{array}{l} \mathrm{C_6H_5NSO_2Cl} \\ \quad\;\, | \\ \quad \mathrm{COCH_3} \end{array} \xrightarrow[\mathrm{NaOH}]{\mathrm{H_2O}} \mathrm{C_6H_5NH_2} + \mathrm{CH_3COO^-} + \mathrm{Cl^-} + \mathrm{SO_4^{--}}$$

Warm alkali splits phenylbenzoylsulfamyl chloride still more slowly, the hydrolysis being arrested at the benzanilide stage.

(Aryl)-arylsulfonylsulfamyl chlorides are split to the corresponding sulfonamides only after a 1-hr. treatment with boiling alkali.

$$\begin{array}{l} \mathrm{ArNSO_2Cl} \\ \quad\, | \\ \quad \mathrm{SO_2Ar'} \end{array} \xrightarrow[\mathrm{NaOH}]{\mathrm{H_2O}} \begin{array}{l} \mathrm{ArN^-} \\ \quad\, | \\ \quad \mathrm{SO_2Ar'} \end{array} + \mathrm{SO_4^{--}} + \mathrm{Cl^-}$$

(*c*) *Acids.* The action of mineral acids on both acyl and arylsulfonyl derivatives is identical with that of alkalies, but is much less rapid.

3. Action of sodium ethoxide or sodium phenoxide

The reaction between sodium alcoholates and dialkylsulfamyl chlorides is one of the general methods for preparing trisubstituted sulfamic acids (137).

$$\mathrm{RR'NSO_2Cl + NaOR'' \rightarrow RR'NSO_3R'' + NaCl}$$

The first product is the true ester of an *N,N*-disubstituted sulfamic acid, but it may rearrange to form the betaine-like salt, $\mathrm{RR'R''NSO_3}$.

4. Action of ammonia and amines

Treatment of dialkylsulfamyl chlorides with ammonia, alkyl amines, or aryl amines yields the corresponding substituted sulfamides (26, 27). The equation for the reaction with ammonia is

$$RR'NSO_2Cl + 2NH_3 \rightarrow RR'NSO_2NH_2 + NH_4Cl$$

The sulfamides resulting from amines are formed in an analogous manner.

The (aryl)-arylsulfonylsulfamyl chlorides, in their overall reaction with ammonia or with primary or secondary amines, appear to act primarily as intermediates for introducing a sulfonyl group between two nitrogen atoms. In this reaction, the original sulfamyl chloride is degraded to the sulfonamide from which it was prepared.

$$\begin{array}{l} ArNSO_2Cl \\ | \\ SO_2Ar' \end{array} + 3R_2NH \rightarrow \begin{array}{l} ArNH \\ | \\ SO_2Ar' \end{array} + R_2NSO_2NR_2 + R_2NH \cdot HCl$$

When R_2NH in the above equation represents a primary or secondary aliphatic amine, the reaction constitutes a general method for the production of *N,N'*-dialkylsulfamides or *N,N,N',N'*-tetraalkylsulfamides (45). On the other hand, when both R groups represent hydrogen, that is, when ammonia itself is used, the reaction produces what is presumably a mixture of sulfamide and trisulfimide.

On aromatic amines, the action of (aryl)-arylsulfonylsulfamyl chlorides is primarily one of oxidation. Thus, with aniline sulfur dioxide is evolved, while the residue consists of aniline hydrochloride, oxidation products of aniline, and the arylsulfonamide from which the original sulfamyl chloride was made (90).

The substitution of an acyl group for the arylsulfonyl group in an organic sulfamyl chloride causes a profound difference in the reaction of the resulting molecule toward ammonia and amines. Undoubtedly, the first action of ammonia and amines on (aryl)-acylsulfamyl chlorides is the replacement of the chlorine. This is followed by ammonolysis or aminolysis of the resulting molecule.

$$\begin{array}{l} ArNSO_2Cl \\ | \\ COR \end{array} + 2NH_3 \rightarrow \begin{array}{l} ArNSO_2NH_2 \\ | \\ COR \end{array} + NH_4Cl$$

$$\begin{array}{l} ArNSO_2NH_2 \\ | \\ COR \end{array} + NH_3 \rightarrow ArNHSO_2NH_2 + RCONH_2$$

Although acetamide, which should be a product of the reaction when phenylacetylsulfamyl chloride is treated with ammonia, has not been isolated, the corresponding *N,N*-dimethylacetamide has been identified

when dimethylamine was employed. The reaction between this class of sulfamyl chlorides and ammonia or amines is an important general method for the preparation of organic sulfamides (4, 7, 90).

The action of ammonia on the only diacylsulfamyl chloride known, *viz.*, the phthalyl derivative, constitutes a method for the preparation of sulfamide. On the other hand, phthalylsulfamyl chloride and amines form stable organic sulfamides of the type $C_6H_4(CO)_2NSO_2NRR'$ (4).

5. *Action of metallic fluorides*

When a dialkylsulfamyl chloride is heated in an indifferent diluent with potassium, sodium, or zinc fluorides, a reaction occurs which merely results in the replacement of the chlorine by a fluorine atom. The substituted sulfamyl fluorides thus prepared are patented for use as insecticides (75).

V. SULFAMIDE

A. PREPARATION

Sulfamide, $NH_2SO_2NH_2$, the diamide of sulfuric acid, was first isolated and characterized by Regnault (106). However, even at the present time, despite a great amount of research, it is difficult to prepare in any considerable quantity. Practically all the methods which have been investigated depend upon the ammonolysis of either sulfuryl chloride or sulfuryl fluoride or of chlorosulfonyl imides. Theoretically the following equation represents a possible reaction between sulfuryl chloride and ammonia:

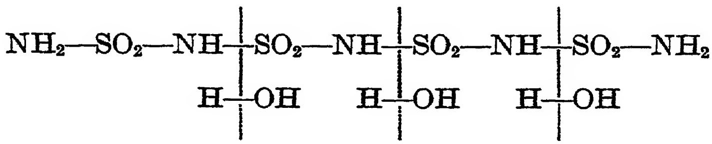
$$SO_2Cl_2 + 4NH_3 \rightarrow SO_2(NH_2)_2 + 2NH_4Cl$$

Unfortunately, there is no indication that the ammonolysis of sulfuryl chloride gives very much sulfamide as a primary product.

The earlier investigators allowed gaseous ammonia to react with an excess of sulfuryl chloride dissolved in an inert solvent such as chloroform, petroleum ether, or carbon tetrachloride. The reaction product was obtained as a gelatinous mass containing large quantities of ammonium chloride (106, 127, 128). For the most part the chloride was tediously removed by precipitation with silver or lead oxides. None of the early workers obtained pure sulfamide, as the melting points reported by them varied between 75°C. and 81°C.

Ruff (111) was the first one to reverse the procedure by keeping ammonia in excess. He removed sulfamide from the reaction mixture by extraction with ethyl acetate and obtained pure sulfamide (m. p. 93°C.). A careful study of the reaction product obtained when ammonia is kept in excess was carried out by Ephraim and Michel (55). They concluded that long-chain compounds of the type $NH_2(SO_2NH)_nSO_2NH_2$ were formed by the

action of sulfuryl chloride upon ammonia in excess and that chains containing four —SO_2— groups were the most common. They pointed out that the very small yield of sulfamide could be ascribed to the only way in which such chain compounds would probably hydrolyze.



Each one of these chains would be able to furnish only one molecule of sulfamide to three molecules of sulfamic acid.

The method for preparing sulfamide from sulfuryl chloride and liquid ammonia was developed by Ephraim and Gurewitsch (53). This method is used today in preference to any other (143). One of the main products of the reaction is thought to be imidodisulfamide, $HN(SO_2NH_2)_2$. The entire liquid ammonia reaction product is dissolved in a weak nitric acid solution and allowed to hydrolyze overnight.

$$HN(SO_2NH_2)_2 + HOH \rightarrow NH_2SO_2NH_2 + NH_2SO_2OH$$

The solution is then evaporated to dryness under reduced pressure and the residue extracted with ethyl acetate. Sulfamide, melting at 93°C., is recovered from the ethyl acetate.

The reaction between sulfuryl fluoride and ammonia is apparently much simpler, leading directly to sulfamide (134). Curiously enough, aqueous ammonia may be used. While sulfuryl fluoride appears to be quite resistant to hydrolysis it does undergo ammonolysis readily. Schlesinger and Shroyer (120) found this procedure to be the most satisfactory one. Wood (143), however, stated that this method could not be recommended, because of the difficulty in handling gaseous sulfuryl fluoride and preparing it in the pure state.

Ammonolysis of aryl chlorosulfonates and *N*-(chlorosulfonyl)acylamides leads to sulfamide, presumably through the intermediate formation of monomolecular and trimolecular sulfimide (4, 5).

A patent (93) has recently been issued describing the preparation of sulfamide directly from sulfur trioxide and ammonia.

$$SO_3 + 2NH_3 \rightarrow SO_2(NH_2)_2 + HOH$$

According to the disclosure and claims the success of this process depends mainly upon two factors: (*1*) Pure dry ammonia must be used and must always be present in large excess, and (*2*) the water which is formed must be removed immediately from the reaction zone. The patent, further, claims that sulfamide may be extracted from the mixture containing am-

monium imidodisulfonate, sulfamic acid, and ammonium sulfate by means of acetone.

B. PHYSICAL AND CHEMICAL PROPERTIES

Sulfamide is a white crystalline solid, which melts sharply at 93°C. (53, 111). It is very stable in air. Although sulfamide is quite soluble in water, its solutions are practically non-conducting. Sulfamide, however, is capable of acting as a dibasic acid, for it forms insoluble salts with the heavy metals, e.g., $SO_2(NHAg)_2$, and soluble salts of similar structure with the alkali metals.

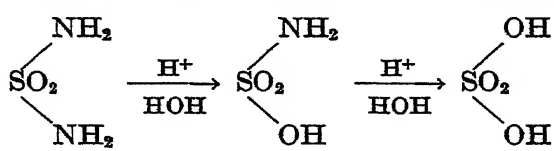
There is a very striking similarity between the structures of sulfamide and urea. An interesting parallelism existing between the formal deammonation products of these compounds is depicted in chart IV. These possible deammonation steps have not all been realized experimentally. It is a well-known fact, however, that biuret is formed when urea is heated; at higher temperatures cyanuric acid is obtained. It has also been demonstrated that triuret is readily deammonated to cyanuric acid (58). Imidodisulfamide and trisulfimide are among the products formed when sulfamide is fused (55, 128). Trisulfimide could also probably be formed by the loss of ammonia from the open-chain trisulfamide (58).

$$NH_2SO_2NHSO_2NHSO_2NH_2 \rightarrow (SO_2NH)_3 + NH_3$$

Sulfamide is stable in neutral, dilute acidic, and dilute alkaline solutions. It readily loses one molecule of ammonia when boiled with strong alkali, to yield the corresponding sulfamate.

$$NH_2SO_2NH_2 + NaOH \rightarrow NH_2SO_2ONa + NH_3$$

When boiled for a long time with hydrochloric acid complete hydrolysis is effected. The stepwise reaction may be represented as follows:



Sulfamide is ammono deliquescent (53, 62, 120) and is very soluble in liquid ammonia. The resulting solution has marked electrical conductivity (61, 62). This is in sharp contrast with its non-electrolytic character in aqueous solution, but again emphasizes its similarity to urea which also acts as an acid in liquid ammonia. Its behavior as an acid in liquid ammonia is in line with its characterization as an aquo-ammonosulfuric acid, just as urea is an aquo-ammonocarbonic acid.

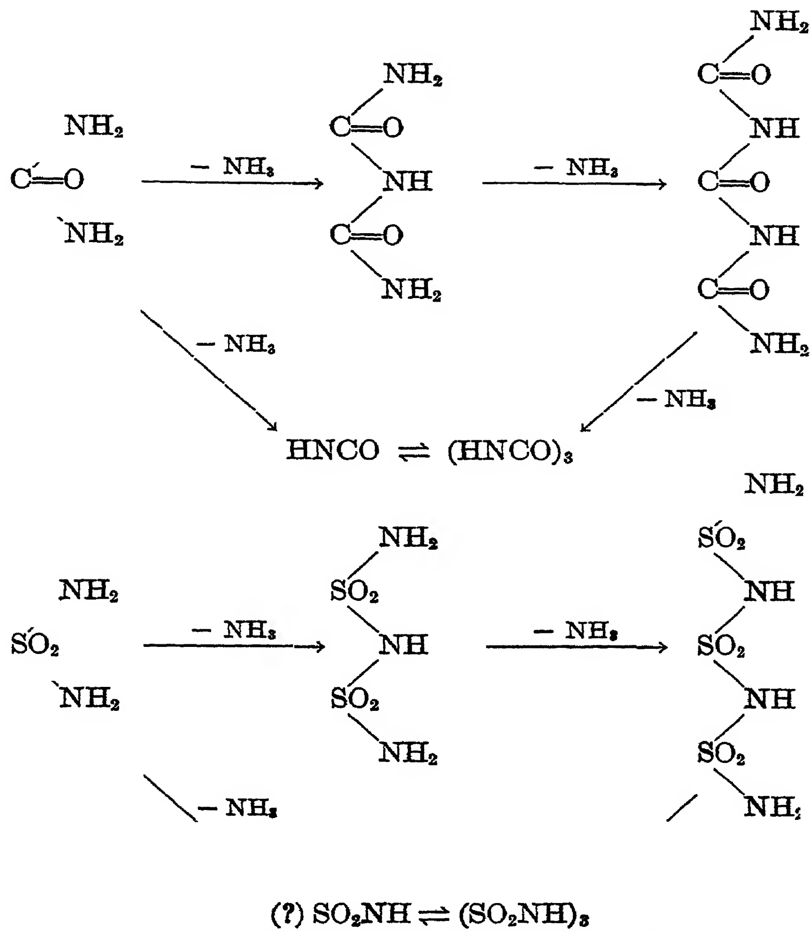


CHART IV. Parallelism existing between formal deammonation products of urea and sulfamide.

C. CHEMICAL REACTIONS OF SULFAMIDE

Although sulfamide dissolves readily in concentrated nitric acid, nitration does not occur unless concentrated sulfuric acid is present, in which case nitrosulfamide precipitates immediately (54).

$$NH_2SO_2NH_2 + HNO_3 \xrightarrow{H_2SO_4} NH_2SO_2NHNO_2 + HOH$$

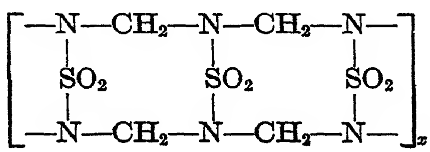
Monochlorosulfamide is formed when an aqueous solution of hypochlorous acid is allowed to react with a concentrated solution of sulfamide (54).

$$NH_2SO_2NH_2 + HOCl \rightarrow NH_2SO_2NHCl + HOH$$

This compound can be obtained in crystalline form (m.p. 66°C.). It is soluble in water, ethanol, ether, and chloroform. In the presence of alkali or ammonia it decomposes and nitrogen is evolved.

$$2NH_2SO_2NHCl + 2NaOH \rightarrow 2NH_2SO_2ONa + 2NaCl + N_2$$

The reaction between sulfamide and formaldehyde has also been studied. This investigation was undoubtedly inspired by the striking similarity between sulfamide and urea. Wood and Battye (143) found that sulfamide and formaldehyde reacted to give a hard water-clear resin. This resin was found to be insoluble in such organic solvents as ethanol, ethyl acetate, toluene, ligroin, and xylene, but slightly soluble in acetone. It softens when placed in warm water and is entirely soluble in boiling water. These authors postulate that the structure of the polymer is similar to that given for the urea–formaldehyde resin, the —CO— group being replaced by the —SO_2— group.



D. METALLIC DERIVATIVES OF SULFAMIDE

Attention has already been called to the marked solubility of sulfamide in liquid ammonia. Both the mono- and di- sodium and potassium salts (63) have been prepared by interaction with the corresponding amides in liquid ammonia. Sulfamide appears to be much more capable of acting as a proton donor, even in aqueous solution, than urea. Insoluble lead (127), mercuric (127), and silver (50, 53, 69, 113, 127) sulfamides, as well as the alkali and alkaline-earth salts (134) and a number of rhodium complexes (87), have been prepared.

VI. THE NITROGEN-SUBSTITUTED SULFAMIDES

A classification of known organic derivatives of sulfamide based upon the number and nature of the nitrogen substituents is proposed here in order to emphasize their diversity. An example of each class is given for illustration.

1. *N,N'*-dialkylsulfamides, $CH_3NHSO_2NHCH_3$ (26, 57)
2. *N,N*-dialkylsulfamides, $(CH_3)_2NSO_2NH_2$ (26, 27)
3. *N,N,N',N'*-tetraalkylsulfamides, $(C_2H_5)_2NSO_2N(C_2H_5)_2$ (45)
4. *N*-monoarylsulfamides, $C_6H_5NHSO_2NH_2$ (7, 90)
5. *N,N'*-diarylsulfamides, $C_6H_5NHSO_2NHC_6H_5$ (90, 126, 142)

6. *N,N′*-dialkyl-*N,N′*-diarylsulfamides, $\left(\begin{matrix} C_6H_5 \\ CH_3 \end{matrix}\!\!>\!N\right)_2 SO_2$ (142)
7. *N,N*-dialkyl-*N′*-arylsulfamides, $(CH_3)_2NSO_2NHC_6H_5$ (7, 90)
8. *N*-monoalkylidene sulfamides, $C_6H_5CH{=}NSO_2NH_2$ (134)
9. *N*-acyl-*N,N′*-diarylsulfamides, $\begin{matrix} CH_3CO \\ C_6H_5 \end{matrix}\!\!>\!NSO_2NHC_6H_5$ (142)
10. *N,N′*-diacyl-*N,N′*-diarylsulfamides, $\left(\begin{matrix} CH_3CO \\ C_6H_5 \end{matrix}\!\!>\!N\right)_2 SO_2$ (142)

A. PREPARATION

Sulfamides are not produced in the reactions between sulfuryl chloride, or dialkylsulfamyl chlorides, and amines unless the basicity of the amines is above a certain rather definite minimum. Thus, aniline treated with sulfuryl chloride yields *N,N′*-diphenylsulfamide without any difficulty, whereas nitroaniline, acetanilide, and phenylammonium chloride treated with the same inorganic reagent undergo ring-chlorination. On the other hand, methylaniline, which is certainly more basic than aniline, is oxidized by sulfuryl chloride to a dark green mass.

However, the sulfamides which should theoretically result from the reaction of sulfuryl chloride with nitroaniline, acetanilide, or methylaniline can be prepared indirectly. Thus, *N,N′*-diphenylsulfamide can be acetylated or methylated to give two of the desired derivatives, and the acetylated product can be nitrated to give the third.

Sulfuryl chloride may be used either directly on amines, or it may be in the form of a derivative before being introduced into the reaction mixture which finally produces the organic sulfamide. The derivatives of sulfuryl chloride which may be employed are phenyl chlorosulfonate, dialkylsulfamyl chlorides, (aryl)-arylsulfonylsulfamyl chlorides, (aryl)-acylsulfamyl chlorides, phthalylsulfamyl chloride, or sulfamide. It should be emphasized that the synthesis of all of these intermediates, save possibly that of sulfamide, requires sulfuryl chloride as an essential reagent.

Since the reaction between the various types of sulfamyl chlorides and ammonia or amines to produce organic sulfamides has already been discussed (page 75), only the methods utilizing phenyl chlorosulfonate, sulfamide, or sulfuryl chloride will be included here.

1. Action of sulfuryl chloride on amines

It has already been emphasized that all types of amines will not react with sulfuryl chloride to produce symmetrically substituted sulfamides.

Apparently it makes little difference which strictly aliphatic amine is employed, be it either primary or secondary in character, for amines with such widely differing organic substituents as methyl, benzyl, dimethyl, and cyclopentamethylene form the corresponding sulfamides. On the other hand, of the aromatic amines, only aniline and *p*-toluidine are known to react smoothly with sulfuryl chloride to give symmetrically substituted sulfamides.

2. Action of phenyl chlorosulfonate on amines

Phenyl chlorosulfonate, $C_6H_5SO_3Cl$, when treated with either mono- or di-alkylamines, first donates the chlorosulfonyl group to the amine with the formation of a sulfamyl chloride, which then immediately reacts with another molecule of the amine to produce the sulfamide. Phenol and hydrogen chloride are by-products of the reaction (45). The equation for the overall reaction is

$$C_6H_5SO_3Cl + 3R_2NH \rightarrow C_6H_5OH + R_2NSO_2NR_2 + R_2NH \cdot HCl$$

The most obvious mechanism would postulate the formation of an ester, a phenyl sulfamate, which would then be aminolyzed to produce the sulfamide. However, Denivelle (45) has shown that phenyl cyclopentamethylenesulfamate, $(CH_2)_5NSO_3C_6H_5$, which is formed from sodium phenoxide and cyclopentamethylenesulfamyl chloride, undergoes no change when refluxed with piperidine. On the other hand, cyclopentamethylenesulfamyl chloride reacts immediately with piperidine to yield dicyclopentamethylenesulfamide.

3. Action of sulfamide on various reagents

The direct replacement of the hydrogens on free sulfamide has not been studied extensively, probably because of the relative difficulty encountered in the preparation of the starting compound. Only three organic sulfamides have been prepared in this manner, *viz.*, benzal (134), tetramethylol (143), and dixanthyl (145). The first two were prepared by the action of benzaldehyde and formaldehyde, respectively, on sulfamide, while the third was isolated from the reaction mixture obtained when sulfamide was treated with xanthydrol.

Arylsulfamides in which the aromatic residue does not contain highly negative substituents like nitro groups behave as acids toward aqueous sodium hydroxide but not toward aqueous sodium bicarbonate, provided the nitrogen carrying the simple aryl radical carries a free hydrogen as well. In other words, this type of sulfamide behaves like an acid toward phenolphthalein. (Should nitro groups be attached to the ring, the resulting sulfamide is easily split by hydrolysis.) Obviously, then, a

sulfamide which dissolves in sodium hydroxide merely because of salt formation can be recovered unchanged by passing carbon dioxide into the solution.

As mentioned previously in this discussion, certain amines will not react directly with sulfuryl chloride to produce sulfamides. However, a number of sulfamides which cannot be made directly can be prepared indirectly. The reactions which lead to these difficultly accessible compounds will now be discussed.

Although methylaniline is certainly more basic than aniline, the former does not yield the corresponding sulfamide derivative when treated with sulfuryl chloride at a low temperature. This rather unexpected failure to produce a sulfamide may be caused by the sensitivity of the ring to oxidation, so that the molecule is disrupted before the amino hydrogen can react, with the simple elimination of hydrogen chloride between the amine and sulfuryl chloride. It is well known that the para hydrogen in methylaniline is quite labile, since, for example, the nitroso group in *N*-nitrosomethylaniline is readily rearranged into the ring. Whatever the explanation may be, *N*,*N'*-dimethyl-*N*,*N'*-diphenylsulfamide can not be prepared directly. However, when *N*,*N'*-diphenylsulfamide is treated with metallic sodium and methyl iodide in methyl alcohol, methylation occurs readily to produce the desired product (142).

$$(C_6H_5NH)_2SO_2 + 2CH_3I + 2Na \xrightarrow{CH_3OH} (C_6H_5NCH_3)_2SO_2 + 2NaI + H_2$$

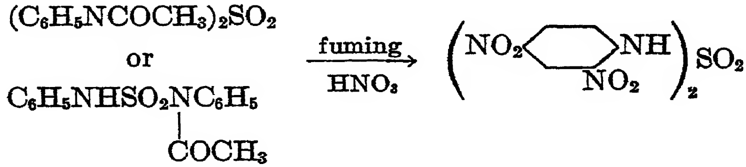
It has already been noted under the discussion of sulfamyl halides (page 72) that sulfuryl chloride exhibits a ring-chlorinating action on acetanilide, and that in order to produce acetylphenylsulfamyl chloride the chlorinating action must be inhibited by employing *N*-sodiumacetanilide. Obviously, then, *N*,*N'*-diphenyl-*N*,*N'*-diacetylsulfamide cannot be prepared directly. If, however, *N*,*N'*-diphenylsulfamide is treated for several hours with acetic anhydride containing a drop of sulfuric acid, the diacetylated product is obtained (142).

$$(C_6H_5NH)_2SO_2 + (CH_3CO)_2O \rightarrow (C_6H_5NCOCH_3)_2SO_2 + H_2O$$

Also under sulfamyl halides (page 75) it was shown that acetylphenylsulfamyl chloride produces *N*,*N'*-diphenylsulfamide when treated with aniline, and not *N*-acetyl-*N*,*N'*-diphenylsulfamide. The latter compound can be prepared merely by decreasing the time during which the acetylating mixture described above is permitted to act on *N*,*N'*-diphenylsulfamide.

Difficulty is again encountered in the reaction between nitroanilines and sulfuryl chloride, and, furthermore, the desired sulfamide cannot be prepared by the most obvious alternative reaction, *viz.*, direct nitration

of N,N'-diphenylsulfamide. If the latter compound is treated directly with fuming nitric acid in concentrated sulfuric acid, a mixture of 3-nitro-4-amino- and 2-nitro-5-amino-benzenesulfonic acids is obtained. These are undoubtedly formed by the rearrangement of the intermediate sulfamic acids. On the other hand, when either the monoacetylated or the diacetylated N,N'-diphenylsulfamide is treated with fuming nitric acid, N,N'-2,4,2',4'-tetranitrodiphenylsulfamide is obtained smoothly (142).



Mention should also be made of the fact that N,N'-dimethylsulfamide, when treated with nitric acid, yields an N,N'-dinitro derivative (142); and that N,N'-diphenylsulfamide forms an N,N'-dinitroso derivative when treated with nitrogen trioxide in ether (57). No other examples of either type of compound have been reported.

VII. Imidodisulfonates

A. PREPARATION

Derivatives of the aquo-ammonosulfuric acid, $HN(SO_3H)_2$, were first prepared in 1834 by Heinrich Rose (107, 108, 109, 110) by the interaction of gaseous ammonia and solid sulfur trioxide. All the methods for preparing imidodisulfonates are in agreement with their designation as aquo-ammonosulfates. Recorded procedures are either ammonolytic or hydrolytic in character, or involve deammonation reactions.

1. Ammonolytic reactions

Depending upon the relative quantities of the reactants, di- and triammonium imidodisulfonates are obtained when gaseous (113), aqueous (19, 20), or liquid (56) ammonia is allowed to react with solid (56) or gaseous (19, 20, 77, 113) sulfur trioxide.

$$4NH_3 + 2SO_3 \rightarrow NH_4N(SO_3NH_4)_2$$

$$3NH_3 + 2SO_3 \rightarrow HN(SO_3NH_4)_2$$

Triammonium imidodisulfonate is obtained by the action of ammonia on chlorosulfonic acid (30, 33). Diammonium imidodisulfonate is obtained when ammonium carbamate is added to pyrosulfuryl chloride, sulfuryl chloride, or chlorosulfonic (89) acid. When molecular amounts of 100 per cent sulfuric acid and urea are mixed at 140°C., the reaction mixture

solidifies to a mass containing diammonium imidodisulfonate and ammonium sulfate. When excess of sulfuric acid or oleum is used, sulfamic acid is obtained (13).

2. Hydrolytic reactions

Imidodisulfonates are produced by careful hydrolysis of nitrilosulfonates (41, 49, 65, 102, 103). If potassium nitrilosulfonate is heated in neutral solution, potassium imidodisulfonate is formed according to the equation:

$$N(SO_3K)_3 + H_2O \rightarrow HN(SO_3K)_2 + KHSO_4$$

This reaction takes place more rapidly in an acid solution.

3. Deammonation reactions

The sulfamates of ammonium, potassium, and lithium (31, 33, 46) have been found to undergo deammonation in accordance with the equation:

$$2NH_2SO_3M \rightarrow NH_3 + HN(SO_3M)_2$$

In the case of the ammonium salt a temperature of 160°C. (31) is required, while for the potassium salt, Divers and Haga (46) give 350°C. as the most suitable temperature.

The fusion of molar quantities of ammonium sulfamate and sulfamic acid leads directly to diammonium imidodisulfonate (13).

$$NH_2SO_3H + NH_2SO_3NH_4 \rightarrow HN(SO_3NH_4)_2$$

B. CHEMICAL PROPERTIES

Free imidodisulfonic acid is apparently not stable in the pure state. An aqueous solution of the free acid was first prepared by Jacquelain (79), who treated the lead salt with hydrogen sulfide. It was also prepared in this way by Berglund (30) and by Divers and Haga (46).

The imidodisulfonate ion may be distinguished from the sulfamate ion by the fact that the former gives a precipitate with barium hydroxide solution or with a solution of barium chloride containing ammonia. This precipitate can be distinguished from barium sulfate, as it is completely soluble in dilute acids.

The imidodisulfonate ion is more stable toward hydrolysis (102, 121) than the nitrilosulfonate ion. However, it does undergo hydrolysis (102, 121) in acid solution to the sulfamate in accordance with the equation:

$$HN(SO_3)_2^{--} + H_2O \rightarrow H_2NSO_3^- + HSO_4^-$$

Although the imido hydrogen is much less acidic than the hydrogen atoms on the sulfonic groups, it is, nevertheless, replaceable by metal

ions. Hence, imidodisulfonic acid forms two series of salts: *viz.*, (*1*) the *basic salts*, those in which all three hydrogen atoms are replaced and (*2*) the *neutral salts*, in which only two hydrogen atoms are replaced. These are represented, respectively, by the type formulae:

$$(1)\ MN(SO_3M)_2 \qquad (2)\ HN(SO_3M)_2$$

With the exception of the dipotassium salt, most of the neutral imidodisulfonates are easily soluble. The basic salts are usually more stable and less soluble in water than the neutral salts (29, 30). The basic salts are readily converted into the corresponding neutral salts by treatment with even weak acids. Conversely, the neutral salts are converted to the basic salts by excess base.

The so-called mercuri-salts form a very interesting group of imidodisulfonates. They are derivatives of an acid with the formula $(HO_3S)_2$-$NHgN(SO_3H)_2$. The parent acid appears to be very strong, but is not very stable. It may be obtained by treating the barium salt with sulfuric acid (30).

C. ALKYL DERIVATIVES OF IMIDODISULFONIC ACID

Traube and Wolfe (136) found that an aqueous solution of tripotassium imidodisulfonate reacts with alkyl halides or dialkyl sulfates to give *N*-alkyl derivatives of dipotassium imidodisulfonate.

$$KN(SO_3K)_2 + RI \rightarrow RN(SO_3K)_2 + KI$$

$$KN(SO_3K)_2 + R_2SO_4 \rightarrow RN(SO_3K)_2 + RSO_4K$$

These reactions give good yields in the case of the methyl and ethyl compounds. However, with other alkyl halides, e.g., propyl or benzyl iodides, the reaction takes place much more slowly and the yields of the alkylated products are smaller. Traube and Wolfe explain these low yields by assuming that the composition of a solution containing one mole of potassium hydroxide and one mole of dipotassium imidodisulfonate is represented by the following equilibrium:

$$HN(SO_3K)_2 + KOH \rightleftarrows KN(SO_3)_2 + H_2O$$

Two possible reactions can, therefore, take place when the alkylating agent is added to this system: (*1*) formation of the *N*-alkylimidodisulfonate from the tripotassium salt and (*2*) reaction of the alkylating agent with potassium hydroxide to give either the alcohol or the olefin.

The alkylimidodisulfonates are stable well-crystallized salts. They are stable in alkaline or neutral solutions, but in acid solution undergo hydrolysis to yield the *N*-alkylsulfamic acids.

$$RN(SO_3K)_2 + H_2O \rightarrow RNHSO_3K + KHSO_4$$

These alkylsulfamic acids can, in turn, be hydrolyzed by prolonged action of mineral acids to the corresponding amine salts.

$$RNHSO_3K + H_2O + H^+ \rightarrow RNH_3^+ + KHSO_4$$

It is interesting to note that this series of reactions provides a method for preparing amines which is analogous to the Gabriel phthalimide synthesis, the former utilizing the potassium derivative of an inorganic imido compound, $KN(SO_3K)_2$, and the latter utilizing the potassium derivative of an organic imide, phthalimide.

VIII. NITRILOSULFONATES

The nitrilosulfonates, derivatives of the aquo-ammonosulfuric acid, $N(SO_3H)_3$, are of interest mainly because they hydrolyze to imidodisulfonates which, in turn, yield sulfamates. The nitrilosulfonates were undoubtedly first prepared by Fremy (64, 65), along with a number of other sulfur–nitrogen compounds. Although he had no idea concerning the structure or nature of these substances, most of his experimental observations were later substantiated. Claus and Koch (39, 40, 41) also investigated the nitrilosulfonates, but considered them to be derivatives of pentavalent nitrogen rather than of ammonia. Despite the fact that much of this early work was later proved to be faulty, these investigators were correct in assuming that the sulfur atoms in the nitrilosulfonates are attached directly to the nitrogen atom. Berglund (33) gave them their name, nitrilosulfonates. Raschig (102, 103) systematized the chemistry of the sulfur–nitrogen compounds, rendering the structure of the nitrilosulfonates still clearer. Divers and Haga (48, 49) devised new methods for preparing some of these compounds and carried out extensive studies on the nature of the nitrite–bisulfite reaction.

A. PREPARATION

The only general method for preparing nitrilosulfonate salts involves the sulfonation of nitrous acid. Raschig (102, 103) represented this stepwise nitridation reaction, as well as the products of the hydrolysis of the intermediate compounds, by the scheme given in chart I (see page 51).

The amount and nature of the products of the nitrite–bisulfite reaction are dependent on the ratio of the reactants. A large excess of bisulfite favors the formation of the nitrilosulfonate, although it has been observed that some nitrilosulfonate is always formed whenever nitrite and bisulfite react, regardless of conditions. Sisler and Audreith (122) studied this reaction and found that the best yields of nitrilosulfonate are obtained (*1*) when the mole ratio of the bisulfite to the nitrite is 4:1 or more and

(*2*) when the bisulfite and nitrite solutions are heated to boiling before mixing.

$$KNO_2 + 4KHSO_3 \rightarrow N(SO_3K)_3 + K_2SO_3 + 2H_2O$$

B. PROPERTIES OF NITRILOSULFONIC ACID

Free nitrilosulfonic acid is not known, since nitrilosulfonates hydrolyze immediately in acid solution. Hydrolysis takes place when a solution of a salt is acidified, or when the neutral solution is heated or allowed to stand for a short time. The initial step in hydrolysis is represented by the following equation (102, 103):

$$N(SO_3)_3^{---} + H_2O \rightarrow HN(SO_3)^{--} + HSO_4^-$$

Imidodisulfonates hydrolyze further to the sulfamate stage, and hence the nitrilosulfonates may be used to prepare sulfamates and sulfamic acid. The hydrolysis of this series of aquo-ammonosulfuric acids is repressed by alkali.

Potassium nitrilosulfonate, $N(SO_3K)_3 \cdot 2H_2O$, is the best characterized derivative of nitrilosulfonic acid (49, 102, 103, 122). It is the most insoluble of the nitrilosulfonates and is indeed more insoluble than potassium perchlorate (49). It loses two molecules of water of crystallization at 100–110°C. and at higher temperatures is decomposed completely. Hydrolysis gradually takes place when the substance stands in air at room temperature, although carefully prepared samples will last about a month before decomposition becomes appreciable. A number of other nitrilosulfonates, both simple (49, 65, 102) and complex (52), have been prepared.

IX. IMIDODISULFAMIDE

A. PREPARATION

A survey of the literature leads to the conclusion that the existence of imidodisulfamide has not yet been demonstrated beyond a question of doubt. Mente (89) claims to have obtained the compound by interaction of chlorosulfonic anhydride and ammonium carbamate. His work has never been substantiated.

Hantzsch and his coworkers (69, 70) claim to have prepared imidodisulfamide starting with the reaction product obtained by passing gaseous ammonia into a solution of sulfuryl chloride in ligroin. After removal of chloride ion from an aqueous solution of this product, the residual material was treated with silver nitrate to precipitate a mixture of the silver salts of various nitrogen–sulfur acids. Silver sulfamide, obtained from this mixture, was decomposed by heating to give the silver salt of trisulfimide, $(AgNSO_2)_3$, which on treatment with hydrochloric acid and hy-

drolysis presumably yielded imidodisulfamide, although Hantzsch first thought a hydrate of trisulfimide had been isolated.

Ephraim and Gurewitsch (53) state that imidodisulfamide is the principal product when sulfuryl chloride is allowed to react with liquid ammonia in excess at low temperatures. At higher temperatures (55), using a solution of ammonia in ligroin, this same reaction presumably causes chain formation to take place with production of compounds of the type $H_2N(SO_2NH)_nSO_2NH_2$.

Most plausible among recorded procedures is the statement that thermal decomposition of sulfamide results in deammonation with formation of some imidodisulfamide (55, 128).

B. PHYSICAL AND CHEMICAL PROPERTIES OF IMIDODISULFAMIDE

Imidodisulfamide is said to be a white crystalline solid melting at 166°C. (70). It is said to be unstable in aqueous solutions, hydrolyzing to form sulfamide and sulfamic acid.

$$HN\begin{array}{l}\diagup SO_2\text{—}NH_2 \\ \diagdown SO_2\text{—}NH_2\end{array} + \begin{array}{c}HOH \\ HOH\end{array} \rightarrow NH_2SO_2NH_2 + NH_2SO_2OH$$

Hantzsch and Stuer (70) present excellent data to show that imidodisulfamide is cleaved instantly by water, even at 0°C. According to them aqueous solutions of imidodisulfamide are strongly acidic because of the sulfamic acid formed by hydrolysis. Ephraim and Gurewitsch (53), however, could not substantiate these findings. Their investigation led them to state that imidodisulfamide is stable for some time, even in acid solution.

In the presence of excess alkali, solutions of imidodisulfamide are apparently stabilized, presumably because of the formation of the corresponding imidodisulfamides. However, these salts have not been isolated. Only the ammonium and silver salts appear to have been characterized. The ammonium salt is formed when pure dry ammonia is passed over the dry imidodisulfamide; its formula is believed to be $NH_3 \cdot HN(SO_2NH_2)_2$. The silver salt, $AgN(SO_2NH_2)_2 \cdot 1.5H_2O$, is slightly soluble in cold but readily soluble in boiling water. It can be recrystallized from hot water without undergoing hydrolysis (70). That the silver is attached to nitrogen seems probable, for treatment with alkali causes no precipitate of silver oxide.

The chemistry of imidodisulfamide needs further investigation in order to clarify the contradictory statements now recorded in the literature.

X. Sulfimide and Trisulfimide

A. METHODS OF PREPARATION

Traube (127) isolated a number of salts from the product he had obtained by the reaction of sulfuryl chloride with ammonia. The analyses of these salts led him to believe that they were derived from the imide of sulfuric acid, $SO_2{=}NH$. He called the salts sulfimides. After many unsuccessful attempts to isolate the free imide he concluded that it could not exist in the solid state.

Later it was established by Hantzsch and Holl (69) that sulfimide did not exist in the simple $SO_2{=}NH$ form in the solution, but that it was present in a trimolecular form, $(SO_2{=}NH)_3$. A ring structure was assigned to this compound, based on its resemblance to cyanuric acid. Hantzsch and Holl (69) even claimed to have isolated "trisulfimide" as a white crystalline substance, melting at 161°C. Four years later, however, Hantzsch and Stuer (70) repeated and extended this investigation and came to the conclusion that the substance reported as free trisulfimide in the earlier work was in reality impure imidodisulfamide. Using extreme precaution to exclude every trace of moisture from their reaction mixture, they isolated a white compound melting at 166°C., which they identified as imidodisulfamide.

All attempts thus far to obtain trisulfimide in the solid state have been unsuccessful. Ammonium acid sulfate is the only product obtained when solutions which should lead to the imide are evaporated. When sulfamide is heated above its melting point and held at a temperature of 170°C. for some time, ammonia is given off. A mixture of trisulfimide and imidodisulfamide is presumably formed.

$$3SO_2(NH_2)_2 \rightarrow (SO_2{=}NH)_3 + 3NH_3$$

$$2SO_2(NH_2)_2 \rightarrow NH(SO_2NH_2)_2 + NH_3$$

It has been suggested recently that trisulfimide is formed as a result of the decomposition of sulfamyl chloride, which is thought to be an intermediate in the reaction between aryl chlorosulfonates and ammonia (45).

$$\begin{array}{l} C_6H_5OSO_2Cl + NH_3 \rightarrow NH_2SO_2Cl + C_6H_5OH \\ \qquad\qquad\qquad\qquad\qquad\quad \downarrow \\ \qquad\qquad\qquad\qquad (SO_2NH)_3 + 3HCl \end{array}$$

During the investigation of the reaction between phosgene and sodium amide by Perret and Perrot, it was noted that derivatives of melanuric acid were present in the reaction products. Because of the structural similarities existing between trisulfimide and melanuric acid, it occurred to the investigators that trisulfimide might be obtainable by the reaction of

sulfuryl chloride with sodium amide (98). By a process of fractional precipitation of their reaction product with silver nitrate appreciable quantities of silver trisulfimide, $(SO_2NAg)_3$, were obtained. It was also demonstrated that no sulfamide is formed in this reaction.

B. DERIVATIVES OF TRISULFIMIDE

It is highly probable that the salts of sulfimide prepared by Traube (127) in 1892 are actually derivatives of trisulfimide. The preparation of the silver salt is given in detail. Large quantities can be prepared by treating the aqueous solution of the product obtained from the reaction between sulfuryl chloride and ammonia with just sufficient silver nitrate to precipitate the chlorine. The precipitated silver chloride is collected and the filtrate, which is acidic, is made ammoniacal. Further addition of silver nitrate causes precipitation of silver trisulfimide, which is only slightly soluble in cold water but fairly soluble in hot water. Its solutions are very stable; the addition of barium nitrate causes no precipitation of sulfate. It is only when the solution, strongly acidified with nitric acid, is boiled that the salt is decomposed and barium sulfate precipitates. The potassium, sodium, and barium salts are also described by Traube (127).

Trimethyltrisulfimide was prepared by Hantzsch and Holl (69) from methyl iodide and silver trisulfimide.

$$3CH_3I + (SO_2{=}NAg)_3 \rightarrow (SO_2NCH_3)_3 + 3AgI$$

It is described as a colorless, odorless, tasteless solid, melting at 121°C. and subliming without decomposition when the temperature is raised above the melting point. The compound is soluble in boiling water and in such organic solvents as ether, chloroform, benzene, and alcohol. Hantzsch and Holl also prepared tribenzoylsulfimide, $(SO_2{=}NCOC_6H_5)_3$, a colorless crystalline compound melting at 112°C.

XI. References

(1) Baeyer: Ann. **127**, 199 (1863).
(2) Bamberger and Hindermann: Ber. **30**, 654 (1897).
(3) Bamberger and Kunz: Ber. **30**, 2274 (1897).
(4) Battegay and Denivelle: Bull. soc. chim. [4] **53**, 1242 (1933).
(5) Battegay and Denivelle: Compt. rend. **194**, 2216 (1932).
(6) Battegay, Denivelle, and Meybeck: Compt. rend. **193**, 1195 (1931).
(7) Battegay and Meybeck: Compt. rend. **194**, 186 (1932).
(8) Baumgarten: Ber. **59**, 1166 (1926).
(9) Baumgarten: Ber. **59**, 1976 (1926).
(10) Baumgarten: Ber. **60**, 1174 (1927).
(11) Baumgarten: Ber. **62**, 820 (1929).
(12) Baumgarten: Ber. **64**, 1502 (1931).
(13) Baumgarten: Ber. **69**, 1929 (1936).

(14) Baumgarten: Ber. **71**, 80 (1938).
(15) Baumgarten: Z. physiol. Chem. **171**, 62 (1927).
(16) Baumgarten: German patent 495,788 (1926); *cf.* Chem. Abstracts **24**, 3517 (1930).
(17) Baumgarten: German patent 519,046 (1931); *cf.* Chem. Zentr. **103**, **I**, 1575 (1932).
(18) Baumgarten: German patent 636,329 (1938); *cf.* Chem. Abstracts **32**, 2959 (1938).
(19) Baumgarten and Krummacher: Ber. **67**, 1257 (1934).
(20) Baumgarten and Krummacher: German patent 598,137 (1934); *cf.* Chem. Abstracts **28**, 5609 (1934).
(21) Baumgarten and Marggraff: Ber. **63**, 1019 (1930).
(22) Baumgarten and Marggraff: Ber. **64**, 301 (1931).
(23) Baumgarten and Marggraff: Ber. **64**, 1582 (1931).
(24) Baumgarten, Marggraff, and Dammann: Z. physiol. Chem. **209**, 145 (1932).
(25) Baumgarten and Olshausen: Ber. **64**, 925 (1931).
(26) Behrend: Ann. **222**, 116 (1894).
(27) Behrend: Ber. **15**, 1610 (1882).
(28) Beilstein and Wiegand: Ber. **16**, 1264 (1883).
(29) Berglund: Ber. **9**, 252 (1876).
(30) Berglund: Bull. soc. chim. [2] **25**, 452 (1876).
(31) Berglund: Bull. soc. chim. [2] **29**, 422 (1878).
(32) Berglund: J. Chem. Soc. [A] **34**, 643 (1878).
(33) Berglund: Lunds Univ. Acta **12**, 12 (1875); **13**, 4 (1876).
(34) Bergstrom: Private communication.
(35) Bucherer and Sonnenberg: J. prakt. Chem. **81**, 1 (1910).
(36) Bühner: Ann. **333**, 283 (1904).
(37) Butler and Audrieth: J. Am. Chem. Soc. **61**, 914 (1939).
(38) Butler, Smith, and Audrieth: Ind. Eng. Chem., Anal. Ed. **10**, 690 (1938).
(39) Claus: Ann. **158**, 52 (1871).
(40) Claus: Ber. **4**, 504 (1871).
(41) Claus and Koch: Ann. **152**, 336 (1869).
(42) Cremer: German patent 654,789 (1937); *cf.* Chem. Abstracts **32**, 3424 (1938).
(43) Cupery: Ind. Eng. Chem. **30**, 627 (1938).
(44) Delepine and Demars: Bull. sci. pharmacol. **30**, 577 (1923).
(45) Denivelle: Bull. soc. chim. [5] **3**, 2143 (1936).
(46) Divers and Haga: J. Chem. Soc. **61**, 943 (1892).
(47) Divers and Haga: J. Chem. Soc. **69**, 1634 (1896).
(48) Divers and Haga: J. Chem. Soc. **77**, 673 (1900).
(49) Divers and Haga: J. Chem. Soc. **79**, 1093 (1901).
(50) Divers and Ogawa: J. Chem. Soc. **77**, 327 (1900); **81**, 504 (1902).
(51) Eitner: Ber. **25**, 461 (1892).
(52) Ephraim and Flügel: Helv. Chim. Acta **7**, 724 (1924).
(53) Ephraim and Gurewitsch: Ber. **43**, 138 (1910).
(54) Ephraim and Lasocki: Ber. **44**, 395 (1911).
(55) Ephraim and Michel: Ber. **42**, 3833 (1909).
(56) Fillinger: Thesis, University of Chicago, 1921.
(56a) Fischer: Ber. **8**, 589 (1875).
(57) Franchimont: Rec. trav. chim. **3**, 417 (1884).
(58) Franklin: *The Nitrogen System of Compounds*, p. 171. Reinhold Publishing Corporation, New York (1935).

(59) Franklin: Reference 58, p. 167.
(60) Franklin: Reference 58, p. 168.
(61) Franklin: Z. physik. Chem. **69**, 272 (1909).
(62) Franklin and Kraus: J. Am. Chem. Soc. **27**, 191 (1905).
(63) Franklin and Stafford: Am. Chem. J. **28**, 83 (1902).
(64) Fremy: Ann. **56**, 315 (1845).
(65) Fremy: Ann. chim. phys. [3] **15**, 408 (1845).
(66) Gordon and Cupery: Ind. Eng. Chem. **31**, 1237 (1939).
(67) Guenther and Holsten: U. S. patent 2,108,886 (1938); *cf.* Chem. Abstracts **32**, 3056 (1938).
(68) Guenther and Holsten: U. S. patent 2,108,887 (1938); *cf.* Chem. Abstracts **32**, 3056 (1938).
(69) Hantzsch and Holl: Ber. **34**, 3430 (1901).
(70) Hantzsch and Stuer: Ber. **38**, 1022 (1905).
(71) Herboth: Arch. Pharm. **262**, 517 (1924).
(72) Hofmann and Biesalski: Ber. **45**, 1394 (1912).
(73) Hunter and Sprung: J. Am. Chem. Soc. **53**, 1443 (1931).
(74) I. G. Farbenindustrie A. G.: British patent 328,032 (1928); *cf.* Chem. Abstracts **24**, 5166 (1930).
(75) I. G. Farbenindustrie A. G.: British patent 457,119 (1936); *cf.* Chem. Abstracts **31**, 3200 (1937).
(76) I. G. Farbenindustrie A. G.: French patent 638,175 (1928); *cf.* Chem. Zentr. **100, II,** 1075 (1929).
(77) I. G. Farbenindustrie A. G.: German patent 562,738 (1931); *cf.* Chem. Abstracts **27**, 1104 (1933).
(78) I. G. Farbenindustrie A. G.: Swiss patent 126,581 (1927); *cf.* Chem. Abstracts **23**, 717 (1929).
(79) Jacquelain: Ann. chim. phys. [3] **8**, 293 (1843).
(80) Junghahn: Ber. **31**, 1234 (1898).
(81) Kirmreuther: Ber. **44**, 3115 (1911).
(82) Kleinberg, Taebel, and Audrieth: Ind. Eng. Chem., Anal. Ed. **11**, 368 (1939).
(83) Krafft and Bourgeois: Ber. **25**, 472 (1892).
(84) Laboratorium "Tasch" A. G.: German patent 484,763 (1927); *cf.* Chem. Abstracts **24**, 1122 (1930).
(85) Linhard: Ann. **535**, 267 (1938).
(86) Mamlock and Wolffenstein: Ber. **34**, 2499 (1901).
(87) Mann: J. Chem. Soc. **1933**, 412.
(88) Marx, Brodersen, and Quaedvlieg: German patent 570,956 (1933); *cf.* Chem. Abstracts **27**, 4246 (1933).
(89) Mente: Ann. **248**, 232 (1888).
(90) Meybeck: Ann. chim. [10] **17**, 129 (1932).
(91) Michaelis and Petow: Ber. **31**, 984 (1898).
(92) Misutsch: Farm. Zhur. **1928**, 310; *cf.* Chem. Zentr. **99, II,** 1129 (1928).
(93) Osswald: U. S. patent 2,117,626 (1938); *cf.* Chem. Abstracts **32**, 5165 (1938).
(94) Paal: Ber. **34**, 2748 (1901).
(95) Paal and Jänicke: Ber. **28**, 3160 (1895).
(96) Paal and Lowitsch: Ber. **30**, 869 (1897).
(97) Pechmann: Ber. **20**, 2539 (1887).
(98) Perret and Perrot: Compt. rend. **199**, 955 (1934).
(99) Piria: Ann. **78**, 31 (1851).

(100) Popelier: Bull. soc. chim. Belg. **35**, 264 (1926).
(101) Ramberg and Kallenberg: Ber. **45**, 1512 (1912).
(102) Raschig: Ann. **241**, 161 (1887).
(103) Raschig: Ber. **20**, 584, 1158 (1887).
(104) Raschig: *Schwefel und Stickstoff Studien*, page 225. Verlag Chemie, Leipzig and Berlin (1924).
(105) Ratti: U. S. patent 1,934,143 (1934); *cf.* Chem. Abstracts **28**, 491 (1934).
(106) Regnault: Ann. chim. phys. [2] **69**, 170 (1838).
(107) Rose: Pogg. Ann. [2] **32**, 81 (1834).
(108) Rose: Pogg. Ann. [2] **33**, 235 (1834).
(109) Rose: Pogg. Ann. [2] **47**, 471 (1839).
(110) Rose: Pogg. Ann. [2] **49**, 183 (1840).
(111) Ruff: Ber. **36**, 2900 (1903).
(112) Sakurai: J. Chem. Soc. **69**, 1654 (1896).
(113) Schaefer and Griessbach: German patent 564,123 (1931); *cf.* Chem. Abstracts **27**, 1101 (1933).
(114) Schirmacher and Schirmacher: U. S. patent 1,750,057 (1930); *cf.* Chem. Abstracts **24**, 2142 (1930).
(115) Schmidt: J. prakt. Chem. **44**, 513 (1891).
(117) Schroeter: German patent 601,105 (1934); *cf.* Chem. Abstracts **28**, 7261 (1934).
(118) Schwarz and Giese: Ber. **67**, 1108 (1934).
(119) Seyewetz and Bloch: Bull. soc. chim. [4] **1**, 320 (1907).
(120) Shroyer: Thesis, University of Chicago, 1928.
(121) Sisler: Thesis, University of Illinois, 1939.
(122) Sisler and Audrieth: J. Am. Chem. Soc. **60**, 1947 (1938).
(123) Smit: Ber. **8**, 1442 (1875).
(124) Spiegel: Ber. **18**, 1479 (1885).
(124a) Sveda: Thesis, University of Illinois, 1939.
(125) Traube: Ber. **23**, 1653 (1890).
(126) Traube: Ber. **24**, 360 (1891).
(127) Traube: Ber. **25**, 2472 (1892).
(128) Traube: Ber. **26**, 607 (1893).
(129) Traube: Ber. **46**, 2513 (1913).
(130) Traube: Z. angew. Chem. **38**, 441 (1925).
(131) Traube and Brehmer: Ber. **52**, 1284 (1919).
(132) Traube, Hoerenz, and Wunderlich: Ber. **52**, 1272 (1919).
(133) Traube, Ohlendorf, and Zander: Ber. **53**, 1477 (1920).
(134) Traube and Reubke: Ber. **56**, 1656 (1923).
(135) Traube and von Drathen: Ber. **51**, 111 (1918).
(136) Traube and Wolff: Ber. **53**, 1493 (1920).
(137) Traube, Zander, and Gaffron: Ber. **57**, 1045 (1924).
(138) Wagner: Ber. **19**, 1157 (1886).
(139) Weil and Moser: Ber. **55**, 732 (1922).
(140) Weil and Wassermann: Ber. **55**, 2533 (1922).
(141) Willcox: Am. Chem. J. **32**, 446 (1904).
(142) Wohl and Koch: Ber. **43**, 3295 (1910).
(143) Wood and Battye: J. Soc. Chem. Ind. **52**, pt. 2, 346T (1933).
(144) Wyler: U. S. patent 2,109,952 (1938); *cf.* Chem. Abstracts **32**, 3424 (1938).
(145) Wood: Nature **136**, 837 (1935).

REACTIONS AND REAGENTS IN LIQUID AMMONIA[1]

CHARLES A. KRAUS

Department of Chemistry, Brown University, Providence, Rhode Island

Received August 3, 1939

I. INTRODUCTION

Aside from water, liquid ammonia is the only solvent concerning which there exists a body of knowledge sufficiently extensive to warrant our speaking of it as having a chemistry of its own. From the chemical point of view, the importance of a solvent medium is determined by the number and variety of the reactions that may be carried out in it and this, in turn, depends upon the number, variety, and convenience of the reagents that the solvent affords. When a given substance reacts with a large number of other substances in the same or in a similar manner, we call it a reagent. Reagents may be conveniently divided into five principal groups: namely, (*1*) acids and (*2*) bases; (*3*) oxidizing and (*4*) reducing agents; and (*5*) reagents that effect the transfer of atoms or groups of atoms from combination with one atom to combination with another. The members of the last group of reagents may be termed "synthetic reagents". In water we have a large number of strong acids and bases as well as strong oxidizing agents, but we have only weak, although numerous, reducing agents and few synthetic reagents.

The rather considerable differences that have been found to exist between the chemistry of solutions in liquid ammonia and in water are, in the main, due to the greater inertness of the H—N bond over that of the H—O bond. Because of the high reactivity of the hydrogen atoms in water, salts of the weaker acids and bases are largely, if not completely, hydrolyzed, and reducing agents stronger than hydrogen react with water with evolution of hydrogen. The chemistry of liquid ammonia solutions is characterized by reagents of high reducing power and by stable solutions of salts of exceedingly weak acids,—if, indeed, certain of these substances may be termed acids at all.

[1] Presented at the Symposium on the Chemistry of Liquid Ammonia Solutions, held at Milwaukee, Wisconsin, September, 1938, under the auspices of the Division of Physical and Inorganic Chemistry and the Division of Organic Chemistry of the American Chemical Society.

II. ACIDS AND BASES

Acids and bases in liquid ammonia are relatively weak when compared with corresponding solutions in water. While the number and variety of acids that may be obtained in liquid ammonia (as ammonium salts) is, if anything, larger than in water, the number of bases is exceedingly limited; only the alkali-metal amides have any considerable solubility in ammonia. The study of acid–base reactions in liquid ammonia has, nevertheless, proved very fruitful, largely owing to the fact that in liquid ammonia solution one may study the salts of many acids which are largely, if not completely, hydrolyzed in water.

Since the acids are much weaker chemical agents in ammonia than they are in water, they may frequently be employed advantageously where side reactions are troublesome in aqueous solutions. Thus, germane may be prepared in yields up to 70 or 80 per cent by treating magnesium germanide with ammonium bromide in liquid ammonia; the corresponding reaction in water gives a yield of only about 20 per cent. The alkali-metal amides, in addition to reacting as bases, also are useful reagents for introducing the amide group into combination with various elements, such as carbon. In reactions of this type, the amides act as synthetic reagents rather than as bases; their reactions will be further discussed below.

III. OXIDIZING AGENTS

Oxidizing agents are less conveniently used in liquid ammonia than in water and are much less powerful. The halogens, as such, are not stable in liquid ammonia solution. Nitric and sulfuric acids form the corresponding ammonium salts and have little or no oxidizing power. Alkali-metal permanganates and chromates are soluble and may be employed as rather weak oxidizing agents. On the other hand, liquid ammonia supplies nitridizing agents which, at times, find useful application.

IV. REDUCING AGENTS

Elementary reductions

The strongest homogeneous reducing agents known to chemistry are the solutions of the alkali and alkaline-earth metals in liquid ammonia. These metals in ammonia solution dissociate into normal positive ions and electrons. The reducing power of these solutions is due to the electrons and is practically identical for solutions of all metals.

The metals in ammonia solution react with many of the more electronegative elements according to the equation:

$$A^n + nM = M_nA^n$$

where A^n is an element of valence n and M is an atom of an alkali metal or an equivalent of an alkaline-earth metal. Thus, sodium and potassium

react with all these elements from chlorine to lead, inclusive, with the exception of nitrogen, carbon, silicon, and, probably, germanium. Excepting the alkali-metal halides, compounds of the type M_nA^n are not appreciably soluble in liquid ammonia, but practically all react with additional atoms of the element A to form highly soluble poly salts. Thus, we have such compounds as Na_2S_x, Na_2Te_2, Na_2Te_x, Na_3Sb_7, and Na_4Pb_9, all of which are readily soluble and all of which are electrolytes in liquid ammonia solution.

Bond reductions

Closely allied to the above reactions are those between the alkali metals and univalent and divalent groups such as triphenylmethyl, dimethyltin, and the like. On treating these groups with an alkali metal, there are obtained salts of which the following are examples: $(C_6H_5)_3CNa$, $(CH_3)_3SnNa$, $(C_6H_5)_3GeNa$, $(C_2H_5)_3SiLi$, $(C_6H_5)_3PbNa$, $(CH_3)_2SnNa_2$, and $(C_6H_5)_2GeNa_2$. These compounds are all salts and are valuable as synthetic reagents. Their reactions will be discussed below.

In a sense, the above reactions involve the splitting of a bond of the type A—A. If the bond is not very stable, fission takes place readily, but as the bond becomes more stable, a point is reached where reaction no longer occurs. Thus, the Sn—Sn bond is readily broken whether the substituent groups be aliphatic or aromatic. The same is true of germanium, but the Si—Si bond is not reduced by the alkali metals in ammonia solution. In the case of carbon, the triarylmethyls are readily reduced by the alkali metals, but corresponding reactions in which aliphatic groups are present do not occur.

Double bonds of the type A=A are readily reduced in the case of tin and germanium and either one or both bonds may be broken, depending upon the amount of metal used. With carbon, the C=C bond is not broken in the case of aliphatic compounds (so far as is known), but one bond is often broken in the case of aromatic compounds. Benzene is stable toward the alkali metals in liquid ammonia, but naphthalene and diphenyl are reduced. Reaction takes place as follows:

$$\begin{array}{c}|\ \ \ |\\ C{=}C\\ |\ \ \ |\end{array} + 2Na = \begin{array}{c}\ \ \ \ |\ \ \ \ |\\ NaC{-}CNa\\ \ \ \ \ |\ \ \ \ |\end{array}$$

Generally, one ion of the resulting salt ammonolyzes according to the equation:

$$\begin{array}{c}\ \ \ \ |\ \ \ \ |\\ NaC{-}CNa\\ \ \ \ \ |\ \ \ \ |\end{array} + NH_3 = \begin{array}{c}\ \ \ \ |\ \ \ \ |\\ NaC{-}CH\\ \ \ \ \ |\ \ \ \ |\end{array} + NaNH_2$$

In some cases, the resulting uni-univalent salt is stable; in others, it ammonolyzes to the hydrocarbon $\mathrm{H}\overset{|}{\underset{|}{\mathrm{C}}}—\overset{|}{\underset{|}{\mathrm{C}}}\mathrm{H}$. Any of these salts may be converted to the corresponding hydrocarbon by the addition of an ammonium salt.

Reduction of salts

When salts of soluble metals are treated with other soluble metals in liquid ammonia, a metathetical ionic reaction occurs, the reaction being controlled by solubility relations. Thus, calcium reduces potassium from its bromide because of the low solubility of calcium bromide. In some cases, metals that have no appreciable solubility in liquid ammonia may reduce alkali metals from their salts because of the great depression of the concentration of the ions due to the reducing metal. Potassium, for example, is reduced from its amide by aluminum because aluminum amide is formed, which, being an amphoteric base, forms potassium ammonoaluminate with excess potassium amide, thus depressing the normal concentration of aluminum ions in the solution. The net result is that electrons, due to aluminum, and potassium ions, due to potassium amide, remain in solution as metal.

When a reducing metal reacts with a salt of another metal which is insoluble in liquid ammonia, the latter metal is, in general, reduced to the elementary state. When the reduced metal is inactive toward an alkali (or alkaline-earth) metal, it remains in the free state. Thus, when a silver salt is treated with alkali metal, metallic silver is precipitated. In other cases, when the precipitated metal is capable of reacting with the alkali metal, as with mercury or lead, for example, the reduced metal forms a compound with the alkali metal. (It should be noted in this connection, that, in many instances, a precipitated metal catalyzes the reaction between the reducing metal and ammonia and the resulting amide reacts further with the originally precipitated metal.)

Reduction of covalent compounds

Turning now to reactions of metals with compounds other than salts, we have several types of reactions:

(*1*) with hydrides, we have the reaction:

$$\mathrm{A—H + M = AM + \tfrac{1}{2}H_2}, \text{ and}$$

(*2*) with other elements, we have the reactions:

$$\mathrm{AX + M = MX + A}$$

$$\mathrm{A + M = MA}$$

Here X is a univalent electronegative element, such as a halogen, and A is a univalent group, such as triphenylmethyl, trimethyltin, or the like.

(*1*) The first reaction might be looked upon as the action of an acid on a metal having a reducing power greater than that of hydrogen. Actually, this relationship is only a formal one; reactions of this type take place with substances which have no acidic properties. Triphenylmethane reacts slowly with potassium to produce potassium triphenylmethide and hydrogen. A better example is the reaction of sodium (or potassium) with monogermane:

$$GeH_4 + Na = NaGeH_3 + \tfrac{1}{2}H_2$$

This reaction takes place readily and is quantitative. The compound sodium germanyl, the analog of sodium methyl, is a true salt and is stable in liquid ammonia at its boiling point; the germanyl ion is not reduced further by sodium. A similar reaction takes place between the alkali metals and arsine, as Warren C. Johnson has shown. Doubtless such reactions would occur between alkali metals and other hydrides, such as silane, phosphine, and the like. The reaction of the alkali metals with certain boranes has recently been studied by H. I. Schlesinger and his coworkers at the University of Chicago, with most interesting results.

(*2*) The reactions of the second type are very common in liquid ammonia solution. They are often complicated by side reactions in which that solvent takes part. The compound AM often ammonolyzes in liquid ammonia according to the equation:

$$AM + NH_3 = AH + MNH_2$$

The resulting amide, in turn, reacts with the original halide according to the equation:

$$AX + MNH_2 = MX + ANH_2$$

The net result of the complete reaction is as follows:

$$2AX + 2M + NH_3 = 2MX + AH + ANH_2$$

When, for example, methyl chloride reacts with sodium in liquid ammonia, methane and methylamine are formed in equivalent quantities. The actual mechanism of the above reaction remains uncertain, but it is well established that an alkali-metal amide is formed as an intermediate reagent. When tetramethyltin is treated with sodium in liquid ammonia, reaction occurs as follows:

$$(CH_3)_4Sn + Na + NH_3 = (CH_3)_3SnNa + CH_4 + NaNH_2$$

Since tetramethyltin is stable toward sodium amide, this substance remains in solution or is precipitated. It is by no means certain, however, that

sodium methyl, $NaCH_3$, is actually formed as an intermediate compound which later ammonolyzes to sodium amide and methane.

The most convenient method of preparing compounds of the type MA is to treat the halides AX with an excess of metal. This avoids the necessity of an initial reduction of the compound to the free group A. The chief difficulty met with is due to the fact that the halides of elements of lower atomic number ammonolyze in liquid ammonia. For this reason, the method is usually inapplicable in the case of silicon and, in some instances, germanium. Sometimes the difficulty can be overcome by dissolving the halide in ethylamine, which has a lower hydrolyzing tendency. In this way, the compound $LiGe(C_2H_5)_3$ has been obtained. This method is not applicable in the case of silicon, but it was found possible to reduce triphenylsilicon chloride to triphenylsilicon in ethylamine solution by means of lithium, owing to the formation of a compound between triphenylsilicon and ethylamine. The precise nature of this compound has not been established.

Any highly halogenated compound reacts with the alkali metals in liquid ammonia, but the tendency of such compounds to ammonolyze is so great that few of the lighter elements can be reduced by this method. Cases in point are the boron and silicon halides. Highly halogenated carbon compounds react with the alkali metals, the halogen being quantitatively converted to halide salt. The method is a very general and convenient one for determining halogens in organic compounds. What happens to the carbon residue, however, remains undetermined. The reactions are complicated, owing to the formation of alkali-metal amide which, in turn, reacts with the halides. Carbon tetrachloride reacts immediately with the alkali metals in liquid ammonia with quantitative formation of sodium chloride, but the residual carbon atom enters into a series of exceedingly complex reactions and these reactions cannot be unravelled until it is known how the alkali-metal amide reacts with carbon tetrachloride. This question will be touched upon below.

The bonds between carbon atoms and the less electronegative elements are, in general, rather stable; thus, the bonds between carbon and germanium and silicon are not reduced by the alkali metals. The carbon–tin and the carbon–lead bonds are both reduced; in the case of tetraaryl tin compounds only one bond is reduced, but in the case of lead, complete reduction may occur. We have no knowledge as to how the bonds between carbon and other metallic elements behave in this respect.

V. SYNTHETIC REAGENTS

When a salt, MA, of a strongly positive ion, M^+, and a weakly negative ion, A^-, is treated with a compound, RX, containing a strongly electro-

negative atom, reaction occurs. The negative atom, X, originally combined with another atom or group of atoms, R, goes into the ionic condition and the residual groups are coupled; thus:

$$MA + RX = MX + RA$$

Reagents of this type (MA) are very common in liquid ammonia solution and they are frequently useful for purposes of synthesis. The reactions are similar to those of the metal alkyl salts, such as LiC_2H_5, and they are used chiefly to combine various elements or groups of elements with carbon or with other similar atoms.

Negative ions containing hydrogen

The alkali-metal amides are typical examples of this class of reagents. They react as follows:

$$MNH_2 + RX = MX + RNH_2$$

where R may be an organic group or a group in which the central atom is other than carbon. With methyl chloride, for example, we have the reaction:

$$KNH_2 + CH_3Cl = KCl + CH_3NH_2$$

With triphenylsilicylchloride, we have

$$(C_6H_5)_3SiCl + KNH_2 = KCl + (C_6H_5)_3SiNH_2$$

These reactions are similar to those of the hydroxides and alcoholates with alkyl halides in water or in alcohol. The alcoholates and the phenolates act similarly in liquid ammonia, with the advantage that hydrolysis of these salts does not occur. Thus we have reactions of the type:

$$C_6H_5OK + C_2H_5Br = KBr + C_6H_5OC_2H_5$$

Similar reactions occur with sulfur and doubtless with selenium and tellurium, although these latter reactions have not been studied.

The reactions are particularly valuable in the case of the elements of the third and fourth groups. Warren C. Johnson, of the University of Chicago, has shown that the compound $NaAsH_2$, the arsenic analog of sodium amide, reacts with alkyl halides as follows:

$$NaAsH_2 + RX = Na + RAsH_2$$

With germanium, we have the reaction:

$$NaGeH_3 + CH_3Cl = NaCl + CH_3GeH_3$$

Substituting a hydrogen atom in methylgermane and treating with an alkyl halide, we have:

$$NaGeH_2 \cdot CH_3 + RCl = RGeH_2 \cdot CH_3 + NaCl$$

In this way, the four hydrogen atoms of germane may successively be substituted by any desired alkyl groups.

With phenyl halides, sodium germanyl reacts in a radically different manner.

$$2NaGeH_3 + 2C_6H_5Br = 2NaBr + 2C_6H_6 + Ge_2H_4$$

Evidently, phenylgermane is unstable; benzene splits off, leaving behind a germane of the ethylenic type. Since the product of the above reaction is a solid, it seems reasonable to assume that the structure in the solid state is complex. It is, however, readily soluble in liquid ammonia. With sodium, it yields the salts $NaGeH_2 \cdot GeH_2Na$ and Na_2GeH_2. The fact that these two salts may be formed indicates that the second germanium bond is reasonably stable. Both bonds are broken down in the presence of an excess of reducing metal. The above reaction requires further study.

In preparing potassium (or sodium) germanyl by the action of the metal on monogermane in liquid ammonia, the reaction takes place quantitatively. In preparing the salts of the partially substituted germanes in a similar manner, the reaction is not quantitative; side reactions occur involving the solvent, as a result of which hydrogen is evolved. Similar side reactions occur in the case of other elements. When triethylsilane is treated with lithium in ethylamine solution, the following reaction takes place quantitatively:

$$(C_2H_5)_3SiH + Li + C_2H_5NH_2 = (C_2H_5)_3Si \cdot NH \cdot C_2H_5 + Li + H_2$$

The lithium merely acts as a catalyst for the reaction; an atom of hydrogen is lost by a molecule of silane and by one of amine and the silicon and nitrogen atoms are coupled.

Similar side reactions occur when certain hydrides are treated with potassium (or sodium) amide. The reaction

$$GeH_4 + KNH_2 = KGeH_3 + NH_3$$

is not quantitative. About 10 per cent of the germane reacts in a complex manner with the amide and ammonia; nitrogen is coupled to germanium, and very large quantities of hydrogen are evolved. In the case of triethylsilane, the reaction is quantitative:

$$2(C_2H_5)_3SiH + KNH_2 = KN[Si(C_2H_5)_3]_2 + 2H_2$$

In working with liquid ammonia solutions, it is necessary always to be on the lookout for side reactions involving ammonia. When chloroform is reacted with potassium amide, reaction proceeds quantitatively according to the equation:

$$HCCl_3 + 4KNH_2 = 3KCl + KCN + 3NH_3$$

It might be expected that with carbon tetrachloride reaction would occur according to the equation

$$CCl_4 + 4KNH_2 = 4KCl + C(NH_2)_4$$

and that the carbon tetraamine would go to some lower deammonation product. Actually, an entirely different reaction occurs. When carbon tetrachloride is treated with potassium amide in liquid ammonia, reaction takes place immediately. Potassium chloride is formed quantitatively, and large quantities of gas, consisting of six parts of nitrogen to one part of hydrogen, are evolved. The solution, originally clear when the reaction is completed, soon becomes opalescent and gradually turns brown; finally, free carbon is precipitated. The amount of carbon precipitated is one-half that originally due to carbon tetrachloride; the remaining carbon must pass off in the form of a volatile compound. The details of this reaction have not been worked out. The reaction well illustrates the need for caution in generalizing in the field of liquid ammonia chemistry.

Substituted negative ions

While reagents of the type MA^nH_{n-1} are very useful, their number is limited. Corresponding compounds in which hydrogen has been substituted by organic groups are more stable as well as more numerous. Practically all the electronegative elements from sulfur to lead, inclusive, form compounds of the type MA^nR_{n-1}, where R is an alkyl or aryl group or, occasionally, hydrogen. In a few instances, the substituted groups (R) may themselves be metallo-organic groups. Thus, we have the reaction:

$$HSiCl_3 + 3NaGe(C_6H_5)_3 = 3NaCl + [(C_6H_5)_3Ge]_3SiH$$

(The corresponding reaction with chloroform does not take place.) The product of this reaction is beautifully crystalline. The hydrogen in the product may be substituted by halogens or other groups. It should be noted that the above reaction cannot be carried out in liquid ammonia, since silicochloroform is ammonolyzed by that solvent. The reagent, sodium triphenylgermanide, however, is easily prepared in liquid ammonia.

Reactions of the type

$$R_3ANa + R'_3AX = NaX + R_3A \cdot AR'_3$$

are readily carried out in the case of tin and germanium and less readily in the case of silicon. The difficulty with silicon compounds is that all the silicon halides are rather readily ammonolyzed by liquid ammonia.

Reactions of the type

$$R_3ANa + R'_3A'X = R_3A \cdot A'R'_3 + NaX$$

take place readily in liquid ammonia and provide a general means for coupling two groups through two different central atoms. Reactions of this type could doubtless be extended profitably to elements of the fifth group and, possibly, to others; they seem not to have been investigated thus far. It is of interest to note that bonds of the type A—A and A—A′ are fairly stable in the case of tin and much more stable in the case of germanium and silicon. The stability increases markedly with decreasing atomic number of the central element. The Sn—Sn bond seems to be somewhat more stable with aryl than with alkyl substituents. This is in contrast to what we find with carbon, where only the hexaaryl ethanes are broken down by reducing agents.

Divalent negative ions

With tin and germanium, it is possible to obtain divalent groups of the type R_2A. They are solids which oxidize readily; they evidently possess a complex structure, since they are quite non-volatile. They react readily with the alkali metals in liquid ammonia, as follows:

$$2R_2A + 2Na = NaAR_2 \cdot AR_2 \cdot Na$$

$$R_2A + 2Na = Na_2AR_2$$

These and other similar compounds are important reagents for building up chains of atoms other than carbon. Thus, treating the above salts with the monohalide R_3AX, we have:

$$NaAR_2 \cdot AR_2Na + 2R_3ACl = 2NaCl + R_3A \cdot R_2A \cdot R_2A \cdot R_3A$$

$$Na_2AR_2 + 2R_3AX = 2NaX + R_3A \cdot R_2A \cdot R_3A$$

Syntheses of this type have so far been confined to tin, germanium, and silicon. There is every reason for expecting that similar reactions would be even more readily applicable to the elements of the fifth and, perhaps, the higher members of the sixth group.

RENAL MECHANISMS CONTROLLING COMPOSITION OF THE BODY FLUIDS[1]

DONALD D. VAN SLYKE

The Hospital of the Rockefeller Institute for Medical Research, New York, New York

Received July 19, 1939

Our cells are bathed in a fluid the volume of which Peters (24) and his collaborators have measured and found to constitute about one-fourth of the body. Evolutionary physiologists like to say that, with regard to the concentration of its salts, the fluid represents the primeval ocean from which our ancestors emerged to venture existence on the dry land. Be that as it may, our cells have maintained a fastidious preference for a surrounding medium of constant and definite composition; and if they are outraged by brusque changes in this medium they refuse to perform their aliquot parts of the body's functions, and we become ill. If the insult is too great, some bloc of cells may so far refuse its duties that the entire organism can no longer carry on, and we die.

The composition of the inner fluid is maintained within necessary limits by a balance between velocities of inflow and outflow of water and solutes. The inflow comes from food and drink and from substances given off from the cells themselves, largely waste products of their chemical activities. The rates of inflow vary tremendously with the digestive and other activities of the body. The outflow is chiefly through the kidneys, and by them it must be so regulated that both the volume and the composition of the inner fluid are kept within limits that permit normal function of the cells and our own enjoyment of life.

The studies of renal behavior which I shall sketch this morning originated in a search for a chemical measure of kidney efficiency. The search was begun in the Rockefeller Hospital some score of years ago by Franklin C. McLean, who had been trained in physiology under Carlson at the University of Chicago, and who has now returned as a member of the faculty. McLean published the first American papers in the field (16, 17), and was then called to China to organize the Peking Union Medical School. After that the problem was continued by others at the Rockefeller Hospital. In its beginning, however, the problem was McLean's.

The studies that followed were made by an unusual succession of men,

[1] The Willard Gibbs Lecture, delivered in Chicago, Illinois, May 20, 1939.

most of whom are now making contributions to chemistry or medicine from their own laboratories or clinics. The work on the chemical factors of the problem was begun with G. E. Cullen, now in Cincinnati, was developed further by Hastings, now at Harvard, and was continued by Sendroy, who is now on one of your own faculties at the University of Chicago. Of the men who worked on the clinical and physiological sides of the problem in the early days were Austin and Stillman, then Salvesen, Lundsgaard, Møller and Kirk, later professors of medicine in Oslo and Copenhagen, Linder, now in Cape Town, and McIntosh, now in Montreal. Of our group your city has taken, not only McLean and Sendroy but also Leiter, Alving, and Benjamin Miller. To these men, and to Alma Hiller and C. P. Rhoads, who are still in the Rockefeller Hospital, are due the studies which we shall review.

First I shall venture to outline a bit of purely chemical work on the enzyme urease, which was a necessary preliminary to our attack on the kidney.

KINETICS OF UREASE

As the excretory product for the study of renal function, urea was the preliminary choice, because its excretion forms a major part of the work of the kidneys. A precise micro method was needed for determination of urea, and Cullen and I decided to use the enzyme urease, which splits urea quantitatively into ammonia and carbon dioxide. To control accurately the action of this enzyme we undertook a study of its kinetics (40, 48). This study led to evidence that the enzyme acts by combining in definite proportions with the urea, which after a definite mean time is thrown off as the hydrolytic products ammonia and carbon dioxide. Data on other hydrolytic enzymes,—invertase, maltase, lactase, emulsin, diastase, and arginase,—indicate that they act by combining with their substrates in a similar manner.

To the theory of alternating combination and decomposition we were led by the following observation: When a given amount of urease acted on solutions of varying urea concentration, the speed of hydrolysis increased with urea concentration up to a certain limit, as shown in figure 1. Beyond that further increases in urea concentration did not make the enzyme work any faster.

Our interpretation of this phenomenon was the following, roughly diagrammed in figure 2. Each urea molecule before it is decomposed first combines with the enzyme. Later, after a time interval, the urea is split and thrown off as its products, ammonia and carbon dioxide. The place of combination on the enzyme molecule is thereafter left vacant until another urea molecule makes contact. Then the decomposition is re-

peated. The time required for a single cycle is the sum of the time taken for another urea molecule to hit the vacated combining point, plus the

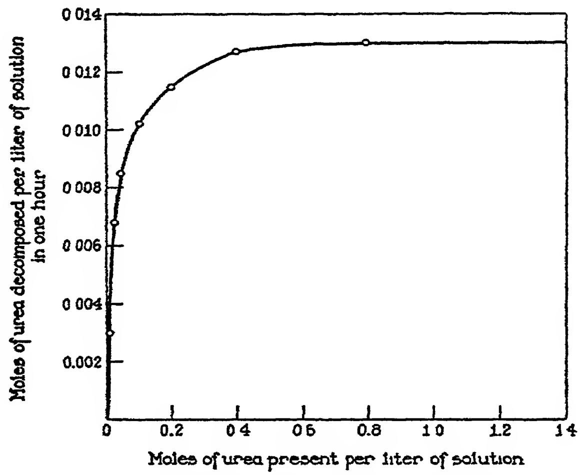


FIG. 1. Effect of urea concentration on rate of urea decomposition by urease (40)

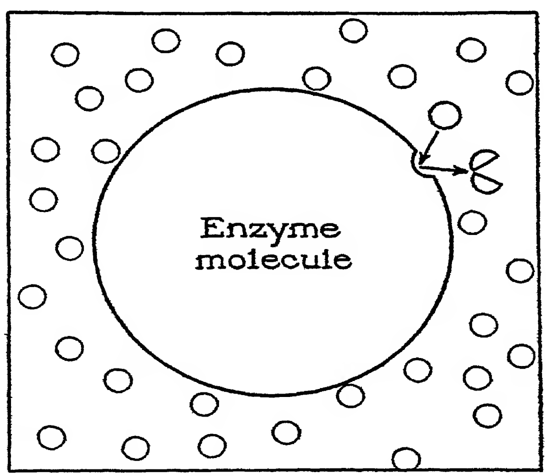


FIG. 2. Diagram of mode of action of the enzyme urease on urea. The urea molecules are represented by the smaller circles. One of them is indicated as having been split into two portions (ammonia and carbon dioxide), which are being ejected from the enzyme. Another urea molecule is indicated as about to take the place of the ejected one in combination with the enzyme.

time the enzyme then takes to split the urea molecule and eject the products. The more abundant the urea molecules are about the enzyme, the

shorter will be the probable path of the next urea molecule to the combining point on the enzyme, and hence the shorter will be the average time interval during which the enzyme is left uncombined and therefore inactive. If the urea concentration is increased enough, this inactive interval may become negligible compared with the interval used for the decomposition; the enzyme works all the time at full speed, because the unused intervals are negligible. Further increase in substrate concentration can cause no further increase in rate of hydrolysis. Under the conditions of the experiment shown in figure 1, full speed was approached when the urea concentration reached 0.6 molar.

To regulate the enzyme action at will one had to learn the effects of conditions on the two separate time intervals, for combination and for decomposition. The reaction was therefore formulated in such a way that the velocity constants governing the two intervals could both be measured.

$$\text{Time required for cycle} = \underbrace{\frac{1}{K_C U}}_{\substack{\text{Interval for com-}\\ \text{bination of enzyme}\\ \text{and urea}}} + \underbrace{\frac{1}{K_D}}_{\substack{\text{Interval for decom-}\\ \text{position of urea}}} \qquad (1)$$

where U = concentration of urea,

K_C = velocity constant of formation of enzyme–urea combination, and

K_D = velocity constant of decomposition of combined urea.

$$-\frac{dU}{dt} = \text{velocity of hydrolysis} \qquad (2)$$

$$= \frac{1}{\text{Time required for cycle}}$$

$$= \frac{1}{\frac{1}{K_C U} + \frac{1}{K_D}}$$

When the urea concentration, U, is large,

$$-\frac{dU}{dt} = K_D \qquad (3)$$

To obtain the time curve of urea decomposition, one integrates the differential equation and obtains:

$\frac{1}{K_C} \ln \frac{A}{U}$	+	$\frac{A-U}{K_D}$	
Time for lowering substrate concentration from its initial value, A, to U	Mean time spent by enzyme molecules uncombined	Mean time spent by enzyme molecules combined with substrate	(4)

The constant K_D, indicating the rate of decomposition of urea after combination with the enzyme, was found by working with urea solutions so concentrated that the inactive periods of the enzyme were negligible, and K_D could be measured by equation 3. After K_D had thus been fixed, the combining constant K_C could be measured by equation 4 in urea solutions so dilute that the time intervals wasted in combining did take up a significant part of the total reaction time.

Thus it was found that K_C, which appears to indicate the relative combining speed of enzyme and urease, is diminished by other crystalloids in solution, either electrolyte or non-electrolyte: their molecules and ions apparently get into the way of the urea molecules and foil part of their attempts to hit the combining point of the enzyme. This effect is important in retarding hydrolysis of the last traces of urea, for it is then that the urea becomes so dilute that the combining speed assumes chief importance in determining the time required to complete the hydrolysis.

The effect of pH, as shown by figure 3, is altogether different on the two reactions. K_C, indicating the speed of combination, increases in a linear way with pH over the entire range from 5 to 9. In contrast, K_D of the decomposing reaction is quickest at pH 7, and slows down as the pH recedes from neutrality in either direction.

Whether the theory of hydrolytic enzyme action outlined above was correct or not, the study of factors affecting the two velocity constants, one predominant at high and the other at low substrate concentration, made possible a quick and exact method of analysis. To get the quickest hydrolysis of urea it became obvious that it is necessary to start the hydrolysis under conditions that favor the decomposing reaction, which predominates in the time consumption in the more concentrated solutions, and to finish under conditions that favor the combining reaction, which consumes increasingly greater proportions of the time as the urea concentration approaches zero. The hydrolysis starts most rapidly in neutral solution and approaches completion most rapidly in alkaline. Buffers are needed to control the pH, but they must be as dilute as possible in order

that their ions and molecules shall not interfere with the combination of enzyme and urea; otherwise the last part of the hydrolysis might be dragged out over a long time.

After these conditions were understood, and a crude but active dry preparation of the urease had been prepared by precipitation of soy bean extracts with acetone, a routine was developed in which urease and urea reacted quantitatively in one minute (39). The mixture was then acidified, and the carbon dioxide of the ammonium carbonate formed from the urea was extracted and measured by the gas pressure it exerted on a manometer. Results accurate to 1 part in 300 could be obtained in rapid routine analyses of 1-cc. samples of blood.

With the enzyme as a tool, we went on to study the action of the kidney.

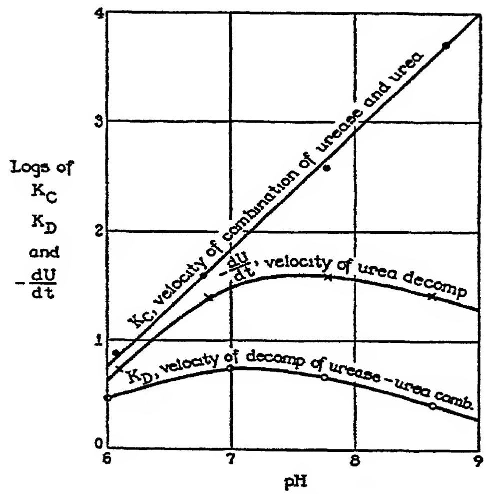


FIG. 3. Effects of pH on the two phases of urease action. Urea concentration, U, = 0.15 molar (47).

UREA CLEARANCE

It had already been shown in the laboratories of Marshall and Addis that the rate of urea excretion in man is proportional to the blood urea concentration. In other words, the amount of urea contained in a given volume of blood is excreted each minute. In normal men excreting abundant volumes of urine, the urea of about 75 cc. of blood is excreted per minute. We therefore say that the normal "blood urea clearance" averages 75 cc. per minute. Expression of the relations among blood concentration, renal excretion, and time, in a single unit capable of visualization (as the volume of blood cleared per minute of excreted substance), was introduced in 1928 by Møller, McIntosh, and Van Slyke (21), and has since been

applied by other authors to the excretion of various substances in renal studies.

Studies of urea clearance with different volume outputs of urine showed that as much as 75 cc. of blood per minute was cleared of urea only when the urine volume was fairly abundant, over 2 cc. per minute (2, 21, 22). Often the urine volume is less than this; in fact the average is about 1 cc. It was found that, as the urine volume fell below 2 cc. per minute, the urea clearance diminished in proportion to the square root of the volume, as exemplified by figure 4. With this relation established[2], when the urine

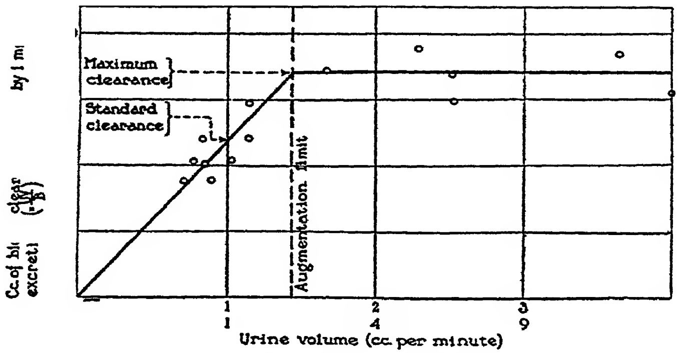


FIG. 4. Relation between the urine volume per minute and urea clearance. U and B indicate urea concentration in urine and blood, V the urine volume in cubic centimeters per minute. When the urine volume falls below the "augmentation limit" of about 2 cc. per minute, the clearance falls parallel with the square root of the volume. (From Møller, McIntosh, and Van Slyke (21).)

volume was low an empirical correction could be made for its effect on the clearance. The probable physiological cause of the effect will be discussed later.

[2] If the urine volume falls below 0.35 cc. per minute Chesley (3) finds that a maximum concentration of urine is reached; with further diminution in urine volume the urea clearance falls in direct proportion to the urine volume instead of to its square root. However, such low volumes are never reached in human urine except under most rigid conditions of water deprivation, or in pathological oliguria.

Dominguez (6), from data of previous authors (1, 21), developed the formula:

$$\text{Urea clearance} = M(1 - e^{-KV})$$

to express the effect of urine volume on the clearance (V = urine volume in cubic centimeters per minute, M = maximum clearance (about 75) approached at high V). This formula, although empirical, gives a curve which approaches the maximal clearance of 75 as an asymptote at high urine volumes, and approximates the square-root rule for volumes between 2 and 0.35 and the linear curve of Chesley for volumes below 0.35. Dominguez' formula has not been applied in clinical or physiological studies, but appears practicable.

Another factor for correction was body size. One could not expect a child to clear of urea as large a volume of blood per minute as a full-sized adult. Comparison of adults and children by McIntosh and Møller showed that the variation was a linear function of the body surface, like the variation in basal metabolic rate (15).

With the clearance thus developed, and corrected for variations in urine volume and body size, it could be applied to studies of renal disease. There it was found to provide a sensitive measure of renal efficiency. In progressive Bright's disease the clearance falls, as the excreting elements are destroyed, until only 3 or 4 cc. of blood are cleared per minute, instead of the usual 75. When the clearance falls to 3 or 4 cc., uremia sets in and death follows (47).

In acute nephritis, however, the clearance may fall almost to the uremic level, and then rise again to normal, with complete recovery (47). The excreting elements are injured, but not irreversibly. An empirical rule was found to hold in acute cases. If recovery is to occur, a rise in the clearance must begin within 4 months after the onset of the acute nephritis Complete recovery may take several months longer, but the upturn of the clearance must begin within 4 months, or as a rule the case will become chronic, and follow a course to fatal uremia. There have been just enough exceptions to prove the rule. I think that two cases in the past 10 years have had their clearances stay at their lowest levels for more than 4 months, and then have staged recoveries; both were children.

EXPLANATION OF CONSTANCY OF CLEARANCE WITH VARYING BLOOD UREA CONCENTRATION

To maintain the observed constant urea clearance of 75 cc. per minute when the blood urea is increased, say tenfold, one of two things must occur. Either the rate of blood flow through the kidney must increase tenfold, with extraction of a constant amount of urea from each cubic centimeter of blood, or else, if the blood flow stays constant, the amount of urea extracted from each unit volume of blood must rise tenfold, so that a constant *proportion* of the blood urea is extracted. Experiments to decide between these two mechanisms required comparison, with respect to urea content, of arterial blood with the venous blood leaving the kidney by the renal vein, in order to find what fraction of the blood urea was removed by the kidney.

We were for a long time stopped by the difficulty of obtaining blood from the renal vein without anesthesia or other unphysiological disturbance. The difficulty was finally overcome by Dr. C. P. Rhoads, who suggested that it might be possible to bring the kidney out to a position under the skin of a dog's back without injury, and suture it there in such a position

that the vein could be tapped by a needle put through the skin, as blood is drawn from the arm vein of a human subject in routine examinations. Rhoads volunteered to attempt the operations and they were completely successful (26). The "explanted" kidneys functioned exactly as they had in their usual positions, and Rhoads was able to needle the renal veins several times a day without inconvenience to the animal.

With this technique we found (45) that an average of one-twelfth of the urea was extracted from the blood of a dog as it flowed through the kidneys. This fraction remained the same, regardless of whether the concentration of urea in the blood was at the usual level of about 0.2 g. per liter, or was raised tenfold by urea feeding.

The blood flow could be calculated by comparing the amount of urea excreted per minute with the amount removed from each liter of blood perfusing the kidneys. For example, if the amount of urea excreted per minute by a dog is found by analyses of arterial and renal blood to be 0.3 of the amount extracted by the kidneys from a liter of blood, the renal blood flow is 0.3 liter per minute.

If in man, as in the dog, the kidneys extract one-twelfth of the urea from the blood that flows through them, the average man's urea clearance of 75 cc. per minute indicates a blood flow of 12 × 75 or 900 cc. per minute. Recent data on human subjects obtained by Chesley and Chesley (4) indicate that 900 cc. per minute is in fact about the mean renal blood flow in man.

The blood supply of the kidney is extraordinarily rich. The entire output of the heart in a resting man is estimated at 4 or 5 liters per minute. Of this about one-fifth goes through the kidneys. Assuming that the kidneys are 1/200 of the body weight, one calculates that, per gram of weight, they receive about thirty five times as much blood as the rest of the body under basal conditions.

PARALLEL VARIATION OF UREA CLEARANCE AND RENAL BLOOD FLOW

We have spoken of the blood urea clearance as though it were a constant value. A certain elasticity is characteristic of the normal kidney, however, and is evidenced by variations in the rate at which it works. A normal man, with abundant urine volume, clearing the urea of an average 75 cc. of blood per minute, may vary the figure between 60 and 100 cc. With dogs, still greater variations occur, and their kidneys can be greatly stimulated by feeding meat diets. As shown in Homer Smith's laboratory (36), if a dog on a cracker meal diet is changed to a meat diet the urea clearance may increase two- or three-fold. The question proposed itself: Is such a physiological increase in the clearance due to acceleration of renal blood flow, or is it due to a more complete extraction of urea from the blood

as it flows through the kidneys? The answer is shown in figure 5. The increase in the urea clearance was found to parallel the increase in blood flow through the dog's kidneys (46). The mean proportion of urea extracted from the renal blood remained at about 8.5 per cent.

FILTRATION–REABSORPTION THEORY OF RENAL EXCRETION

All the results outlined fit the filtration–reabsorption theory of urine excretion. I shall pause to outline this theory, for it assists one to tie together the facts just discussed and to fit into the picture others to be brought out later.

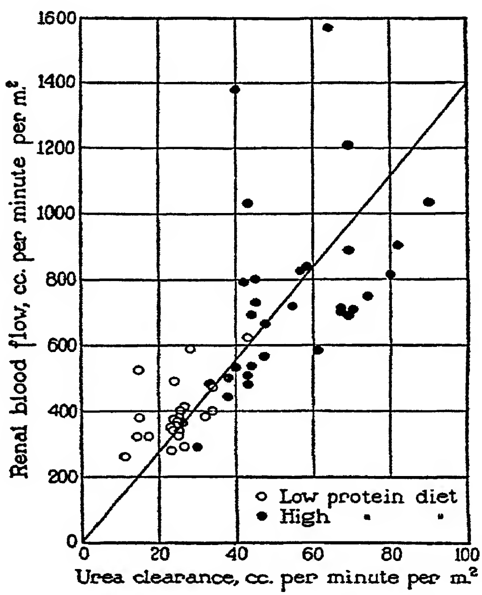


FIG. 5. Parallelism between urea clearance and renal blood flow in dogs. (From Van Slyke, Rhoads, Hiller, and Alving (46).)

The physiologist, Ludwig, realized that the glomerulus was adapted to filtration, and proposed the first filtration–reabsorption theory. He suggested that as blood passed through the capillary loops in the glomerulus (figure 6) some of the water with its dissolved non-colloid solutes was filtered out into the globular space outside the capillary tuft. The filtrate then passed down the tubule, where it was conceived to be concentrated by reabsorption of water.

When the chemistry of blood and urine became better known, however, it became obvious that simple reabsorption of water could not explain the differences between blood and urine concentrations of different solutes.

For example, the ratio, concentration in urine: concentration in blood plasma, averages about 50 for urea and for glucose only about 0.2 (41). For such differences Cushny (5) advanced the hypothesis that the tubules reabsorb different constituents of the glomerular filtrate with varying degrees of completeness, and that in general the tubules reabsorb those substances which the body needs to retain, and refuse more or less completely to reabsorb substances of which there is a surplus, or which are

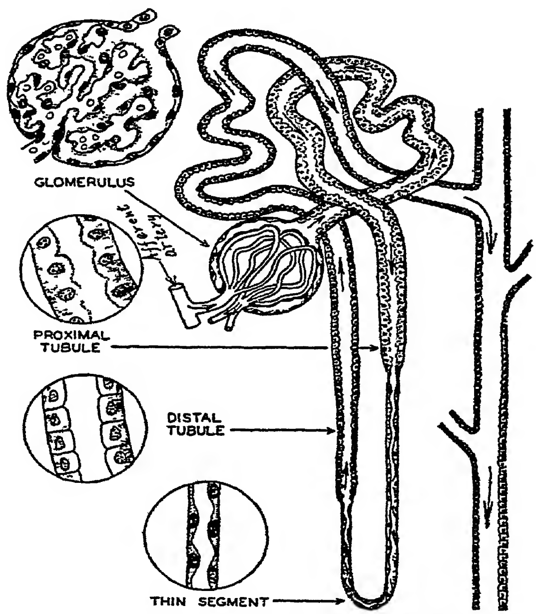


FIG. 6. Diagram of a nephron (from Homer Smith (36)). Blood enters the glomerulus from the afferent artery, flows through the capillary loops of the glomerular tuft, and emerges by the efferent vessels. From these the blood circulates about the tubules in a network of vessels not shown in the diagram. The fluid filtered from the blood in the glomerulus passes down the proximal and then the distal tubule (in the direction shown by the arrows), and has its composition altered by tubular reabsorption of water and some of the solutes, so that it enters the collecting tubule as elaborated urine.

waste products to be gotten rid of. This theory has been brilliantly confirmed by recent work of several laboratories, but in particular those of A. N. Richards (27, 28, 31, 32), Smith and Shannon (34, 35, 36, 37, 38), and Marshall (12, 13).

Richards has tapped single glomeruli of the frog's kidney with microscopic glass pipets, and with his collaborators, Wearn, Hayman, Bordley, Walker, and others (31), has developed ultramicro methods for analysis of

samples of 1 or 2 cmm. of the glomerular filtrate thus obtained. They have found that it is in fact a true filtrate, with urea, uric acid, chloride, creatinine, glucose, and pH in approximately the same concentrations found in the plasma. That the mammalian glomerulus filters in the same manner as demonstrated for the frog's glomerulus has not been proved in the same direct way, but the similar structure makes a similar filtration probable, and all the facts accumulated by exact methods are in harmony with the filtration mechanism for the mammals.

In accord with this mechanism our results with urea may be explained as follows: As the blood flows through the glomerulus (figure 6) a fairly constant fraction of the water is filtered out, with its dissolved urea. The limit of the fraction of plasma water filtered appears to be set chiefly by the balance between the blood pressure within the glomerular capillaries, which forces the filtration, and the opposing osmotic suction of the plasma proteins, which tends to keep water from leaving the plasma. As the blood flows from the afferent artery through the glomerular capillaries the blood pressure presumably decreases, and the osmotic suction of the plasma proteins rises as they become more concentrated from loss of filtered water. To judge from the capillary pressure measurements of Landis (11) and the known osmotic activities of the proteins, it appears probable that the two forces may approach equilibrium at about 25 mm. of mercury pressure before the end of the capillary is reached. In this manner a fairly constant proportion of the plasma water may be filtered.

PROPORTION OF PLASMA WATER FILTERED

In testing the conception of glomerular filtration the question arose: What fraction of the plasma water is filtered out as the blood passes through the glomeruli? If all the urea that is filtered were excreted, one could estimate the fraction of blood water filtered as equal to the fraction of urea removed from the renal blood. However, there is reason to believe that some of the filtered urea diffuses back into the blood with the water reabsorbed from the tubules. For study of the filtered fraction an excretory substance was needed which would not be reabsorbed. It was provided by the anatomist, Gersh (7). He found that when ferrocyanide was excreted by rabbits it could be made visible by precipitation as Prussian blue, and could be seen in the glomerular filtrate and the tubular fluid. However, none appeared in the cells that line the tubules. It appeared that, unlike urea, ferrocyanide passes down the tubules without at all diffusing back into the blood. That the kidney of the dog behaves in the same way towards ferrocyanide was shown later by Gersh and Benjamin Miller, whom Gersh kindly invited to his laboratory, which was then at the Johns Hopkins University, for the necessary experiments.

There were two other substances which, because their clearances were known to be higher than the urea clearance, seemed perhaps able to escape reabsorption in the tubules. One of these was creatinine, which had been suggested by Rehberg (25) as a non-reabsorbed substance. The other was the complex carbohydrate inulin, which Richards, and Smith and Shannon, had found to be excreted by dogs with the same clearance as creatinine.

Consequently we injected ferrocyanide, creatinine, and inulin, dissolved together, into the veins of dogs and observed the proportions that were extracted from the blood plasma by the kidneys (42, 43). For these observations the dogs prepared by Rhoads (26) with explanted kidneys were used, and the percentage extraction was measured by the decrease in plasma concentration of each substance as the blood passed through the kidneys. Such a small proportion of the filtered water escapes reabsorption that the volume of the plasma is practically unchanged by passing

TABLE 1

Extraction of ferrocyanide, creatinine, inulin, and urea from blood plasma by the dog kidney

MATERIAL	NUMBER OF DETERMINATIONS	MEAN EXTRACTION PERCENTAGE	STANDARD DEVIATION
Ferrocyanide	25	18.8	5.5
Creatinine	36	19.9	3.8
Inulin	21	22.3	7.9
Urea	37	8.3	2.0

through the kidneys, and the fall in concentration of dissolved substances in the plasma could be taken as an approximate measure of the proportion removed by the kidneys.

The results (see table 1) showed that all three substances, ferrocyanide, inulin, and creatinine, were extracted from the plasma to the same extent,—about 20 per cent of each on the average was removed as the blood passed through the kidneys.

We believe that the most probable interpretation of these results is that about 20 per cent of the plasma water and its dissolved crystalloids are filtered out in the dog's glomeruli.

PASSIVE REABSORPTION OF UREA FROM TUBULES

That the kidney removes urea from the blood less completely than creatinine or ferrocyanide seems due to back-diffusion of part of the filtered urea into the blood of the tubular network as the filtrate passes down the tubules. That the kidney permits part of a useless excretory product to

diffuse back in this manner must be attributed to an imperfect development of the organ's function. It is what Metchnikoff called one of the anomalies of nature. The kidney would be a more perfect organ if the tubular cells were entirely impermeable to urea, so that all that is filtered passed out in the urine, as in the dog is the case with ferrocyanide and inulin and creatinine. Comparison of the extraction percentages of these substances with that of urea, however, indicates that, even when urine volumes are large enough for maximal urea clearance, about 40 per cent of the filtered urea diffuses back into the blood[3] and has to be brought around to the kidney and filtered again before it is finally gotten rid of (42, 43).

The back-diffusion of urea appears to occur in two phases: a first phase which accompanies reabsorption of the first 90 per cent of the water of the glomerular filtrate, and a second phase which accompanies reabsorption of whatever part of the last 2 per cent of the filtered water is reabsorbed. The absorption of the first 90 per cent of the 120 cc. of water filtered per minute one may call, with Homer Smith, "obligative reabsorption," since it appears imperative in the normal kidney, which seldom lets the urine flow rise above 12 cc. per minute, even under the pressure of water drinking or diuretics. The backflow of this reabsorbed water into the tubular blood must be rapid, and it sweeps with it about 40 per cent of the filtered urea. Further reabsorption of water does not appear to take any more urea with it, however, until the urine volume is contracted to about 2 cc. per minute, since the urea clearance shows no definite change with urine volume over the range of the latter from 12 cc. to 2 cc. per minute (figure 4, horizontal part of the curve). However, when contraction of the urine volume extends below 2 cc. per minute the ratio of urea concentration in the distal tubular lumen to urea concentration in the blood goes above 30, and along this steep osmotic gradient urea begins again to diffuse back into the blood in amounts which increase rapidly with the degree to which the urine is further concentrated (figure 4, sloping part of the curve).

The two phases of urea reabsorption appear to be due, as pointed out by Smith and Shannon (36), to different processes. The second phase, because of the great concentration gradient between the filtrate and the blood, seems attributable to diffusion pressure breaking through the resistance of

[3] The data of table 1, showing only 8.3 per cent decrease in plasma urea concentration as the blood passes the kidneys, compared with 20 per cent for ferrocyanide, inulin, and creatinine concentration, would seem to indicate that not 40 but 60 per cent of the filtered urea gets reabsorbed. However, part of the urea extraction indicated by plasma analysis is masked by quick transfer of urea from red blood cells to plasma when water is reabsorbed in the tubules. When correction is made for this transfer, reabsorbed urea is estimated at 40 per cent (43). The transfer does not occur with creatinine, inulin, or ferrocyanide (43).

the tubular cells. The first phase, however, occurs when the concentration gradient is relatively small, and appears to be in part due to some other contributing force, such as the mechanical sweep of the back-drawn water. However, in both phases urea moves from higher concentrations in the filtrate to lower ones in the blood; there is no active propulsion at any time against a urea concentration gradient, and hence reabsorption of urea in both phases may be termed *passive* insofar as the activity of the tubular cells is concerned. Uric acid appears to be partly reabsorbed by similar passive diffusion.

ACTIVE REABSORPTION BY THE TUBULES

In contrast to urea and uric acid, glucose can be actively pushed by the tubular cells back into the blood, even *against the concentration gradient*, for the process continues after the glucose in the tubular lumen becomes less concentrated than in the blood. Such absorption is called *active*. Besides glucose, water, sodium, potassium, chloride, and other blood crystalloids of which certain amounts must be maintained in the body, are taken back by active, selective reabsorption from the tubular lumina when necessary to prevent the body's supply from falling below the physiological optima. For such substances the cells of the tubular walls act as force pumps, driving the substances back into the blood, when necessary against concentration gradients. Richards and his collaborators (30, 32), by ultra-analyses of fluids drawn from successive segments of amphibian tubules, have mapped the parts where different filtrate substances are reabsorbed.

KINETICS OF ACTIVE REABSORPTION. SIMILARITY TO ENZYME KINETICS

In the conception of this active function of the tubules an important step has been made by Shannon and Fisher (34, 35). By increasing the concentration of glucose in the glomerular filtrate Shannon and Fisher (35) found that the rate of reabsorption could be increased up to a certain limit. Beyond this maximum further increase of glucose in the tubular lumen could not make the sugar pass back into the blood any faster, and it began to escape into the urine. With regard to glucose reabsorption the tubular cells were then acting at full speed, like the urease in figure 1 when the urea concentration exceeded 0.6 molar. Shannon and Fisher's quantitatively formulated explanation was that the glucose combines reversibly in the cells with a hypothetical carrier substance, from which the sugar is later freed to pass into the blood on the other side of the cell. This hypothetical process is analogous to that by which urease and other hydrolytic enzymes appear to combine with their substrates, and after an interval to eject them, in this case as hydrolytic products. Shannon's curves of the

process of glucose reabsorption are of the same character as the curve of urease action in figure 1.

TUBULAR EXCRETION

Besides reabsorbing, the tubular cells can pass materials in the opposite direction, from the blood into the tubular lumen. This process is tubular excretion. When certain substances are injected into the blood, they are almost completely extracted from the plasma that passes through the kidneys. Such is the case with an organic iodine compound called "diodrast" and with the dye phenol red. Marshall (12) has shown that it would not be possible to filter more than a slight fraction of this dye, because 80 to 90 per cent circulates loosely combined with the proteins in the plasma and is non-filterable. Yet the kidneys remove 60 per cent, 70 per cent, or more of it from the plasma. The only way in which this could occur is by an excretory process which is active, like the active reabsorption of glucose, and it appears possible only in the tubules. The tubular cells thus appear to act as two-way carriers, passing water, glucose, and certain other substances against osmotic gradients from the tubular fluid back to the blood, and passing other substances, such as diodrast and phenol red, from the blood into the tubular lumen. Tubular excretion does not ordinarily seem to play an important rôle in maintaining the normal composition of the body fluids of mammals. It is a reserve function, that can come into play in unusual emergencies, as when such a foreign substance as phenol red or diodrast enters the circulation. In some salt water teleosts, such as the toadfish, however, it is the only mode of renal excretion: these fish have lost their glomeruli (8, 37), and perform all their renal excretion by extrusion of substances from the tubular walls (13, 36). They have developed this mode of excretion, while the mammals and many lower animals have developed glomerular filtration, the tubules becoming specialized chiefly in reabsorption of the substances that must be retained. Shannon (34) has shown that his theory for active tubular reabsorption by combination and dissociation applies equally well to active tubular excretion.

COMPARISON OF UREA CLEARANCE WITH DIODRAST CLEARANCE IN MAN

If any constituent is completely removed from the renal blood and excreted by the kidneys, the "clearance" of that substance will be equal to the total volume of blood flowing through the kidneys per minute. There is evidence that tubular excretion enables the normal human kidney to extract the organic iodine compound, diodrast, completely, or almost completely, from the renal blood. Therefore the diodrast clearance has been used by Smith, Goldring, and their collaborators (2, 36, 37, 38) as a measure of the renal blood flow in man. Urea clearances were measured

at the same time. The renal blood flow was caused to vary by such drugs as adrenaline and theophylline. It was found that the urea clearance was unaffected by variations in estimated renal blood flow, when the latter was calculated as equal to the diodrast clearance.

This experimental result is the opposite of that obtained in our laboratory by comparing urea clearance and measured blood flow changes in dogs, when the blood flow variations were produced by adding or withholding meat in the diet. In our dog experiments we found that the urea clearance was not independent of the renal blood flow, but was directly proportional to the flow, the percentage of urea extracted from the renal blood remaining constant despite gross variations in blood flow and clearance. From the constant extraction percentage, it would appear that in our experiments the renal blood probably perfused the glomeruli at a constant pressure, even when the renal blood flow varied, and that the most probable mechanism for varying the flow without the pressure was by opening and closing varying proportions of the glomerular capillaries. Such an opening and closing had been seen by Richards (29) in the capillaries of the frog's glomerulus. In contrast, the results of Smith, Goldring, and their collaborators are most readily explainable on the assumption that the renal blood flow is controlled, not by the opening and closing of capillaries, but by the contraction and expansion of the efferent artery leading from the glomerulus. In such control the same arterial contraction that slowed the blood flow would produce a back-pressure, which would increase the proportion of plasma water and solutes filtered in the glomerulus; as much filtrate per minute might thus be filtered with a slow blood flow as with a rapid one. Smith in fact, concludes (38): "The renal blood flow in man is apparently controlled by the efferent arterial tone."

The diametrical difference between our results with the dog and those of Smith and Goldring with man, concerning the relation of urea clearance to renal blood flow, may be due (*1*) to the difference in technique of estimating renal blood flow, or (*2*) to an actual difference between the mechanisms of the human and canine kidney,[4] or (*3*) to the fact that the renal blood flow

[4] The kidneys of the dog and man do not behave entirely alike with regard to every function. While it appears that in both species the inulin and creatinine clearances (taken without administration of creatinine in the case of man) (20) are equal to the glomerular filtrate, ferrocyanide shows an altogether different behavior in man. Whereas in the kidney of the dog or rabbit both histochemical and functional studies indicate that ferrocyanide is excreted without any reabsorption by the tubules, in man Miller and Winkler (19) found that ferrocyanide was reabsorbed to about the same extent as urea, *viz.*, 40 per cent or more. The cells of man appear to be less resistant to penetration by ferrocyanide than those of either animal. One sign of this is the failure of the tubules of man to prevent back-diffusion of filtered ferrocyanide. Another is that injected ferrocyanide is of the order of tenfold as toxic for man as for the dog; it appears to penetrate the cells of the tissues as well as of the tubules in man more readily than in the dog.

variations in our dogs were induced by diet, while the variations in Smith and Goldring's human subjects were induced by other stimuli, *viz.*, drugs or water diuresis. It may be that under the same stimuli the dog would show the same behavior noted by Smith and Goldring in man. Whether the dog will do so, is one of the questions not yet answered. A positive answer would indicate that the dog under different stimuli could use at least two different vascular changes to vary his renal blood flow, and would lead one to anticipate a similar versatility in man.

MAINTENANCE OF BODY NEUTRALITY BY REGULATION OF BICARBONATE EXCRETION

Besides excreting waste products, the kidney must maintain the neutrality of the body, which produces usually acid, but sometimes alkali, in excess. The regulation by excretion of acid or alkali at need is so sensitive that the pH of the blood seldom varies beyond the range 7.30 to 7.45.

TABLE 2

Estimated ordinary 24-hr. filtration and excretion of bicarbonate, free carbonic acid, and phosphate by man

(From Sendroy, Seelig, and Van Slyke (33))

SUBSTANCE	GLOMERULAR FILTRATE, 170 LITERS AT pH 7.4		URINE, 1.5 LITERS AT pH 6.1	
	Concentration	Total amount	Concentration	Total amount
	millimoles per liter	*millimoles*	*millimoles per liter*	*millimoles*
HCO_3	25	4200	2	3
H_2CO_3	1.3	200	2	3
PO_4	1	170	43	65

The manner in which the kidney produces at will acid or alkaline urine may be deduced from experiments by Sendroy and Seelig (33), showing that bicarbonate is actively reabsorbed, although free carbonic acid is not. Bicarbonate concentration in the urine was found to vary from four times to one-twentieth of the bicarbonate concentration in the blood plasma. The highest bicarbonate concentrations were in the most alkaline (pH 7.5) urines, and the lowest in the most acid (pH 4.9) urines. *Bicarbonate can be so completely reabsorbed that in the more acid urines practically none at all is present.* As seen from table 2, of the bicarbonate filtered only about 1/200 ordinarily appears in the urine, and in the most acid urines as little as 1/2000 may appear. Bicarbonate is the chief form of reserve buffer alkali in the body fluids, and it is by the kidney's power to reabsorb it from the glomerular filtrate that loss is prevented.

Free carbonic acid, on the other hand, was found from one to five times as concentrated in the urine as in the blood, but never less concentrated.

Phosphate is the only other important blood buffer that appears to be filterable. About three-quarters of that filtered appears to be reabsorbed. The effect of its reabsorption on preservation of the phosphate content of the body fluids is vital, but the effect on preserving neutrality appears negligible compared with bicarbonate reabsorption.

Of the urinary constituents causing the acid reaction that is ordinarily present, *viz.*, pH 6 $\pm$ 1, acid phosphate, as shown by Henderson and Palmer (9), is the most important. The large amount of phosphate present, compared with the HCO_3^- and H_2CO_3, appears due to the immensely less complete reabsorption of PO_4.

It does not appear probable that reabsorption of alkaline phosphate and excretion of the unreabsorbed acid phosphate into the urine account for the acidity of the latter. If the fall of urinary pH below blood pH were attained by selective reabsorption of HPO_4^{--}, phosphate excretion would be expected to diminish as the urine became more acid. On the contrary, in conditions of acidosis, where the urinary pH is low, the excretion of total phosphate is not smaller, but greater, than usual.

One is led to the conclusion that the acid phosphate in the urine is chiefly formed by reaction in the tubules of alkaline phosphate with free carbonic acid. Phosphates and carbonates are in equilibrium according to the equation

$$\frac{HPO_4^{--}}{H_2PO_4^-} = 0.2\frac{HCO_3^-}{H_2CO_3} \tag{5}$$

As bicarbonate is withdrawn from the tubular fluid, the ratio $HCO_3^- : H_2CO_3$ is decreased, and in consequence reaction with alkaline phosphate occurs,

$$H_2CO_3 + HPO_4^{--} = HCO_3^- + H_2PO_4^-$$

with lowering of the ratio, $HPO_4^{--} : H_2PO_4^-$, in accordance with equation 5.

It appears accordingly that, although the acidity of the urine is due chiefly to its acid phosphate, the presence of the phosphate in acid rather than alkaline form is due to the manner in which bicarbonate is reabsorbed in the tubules, leaving a relative excess of free carbonic acid to react with the alkaline phosphate. Continued reabsorption, active and passive, removes nearly all of both the bicarbonate and free carbonic acid, but leaves the acid phosphate.

There are times when more alkali than acid is produced in the combustion of foods, or is absorbed as administered sodium bicarbonate, and the kidney is then called upon to get rid of alkali instead of acid. It appears to do this simply by reabsorbing the bicarbonate somewhat less completely

than usually, so that from a fraction of a gram to 2 or 3 g. per hour is let escape into the urine. Since glomerular filtration is calculated to remove 12 to 15 g. of bicarbonate per hour from the blood, the greatest observed bicarbonate excretions can be easily attained merely by decreasing the reabsorption.

MAINTENANCE OF BODY NEUTRALITY BY AMMONIA EXCRETION

To prevent acidification of the body, the kidney has, besides bicarbonate reabsorption, another mechanism in the ability to form ammonia. Nash and Benedict (23) showed that the urinary ammonia is formed in the kidney. Its precursor is presumably either the urea or the amino acids brought to the kidney by the blood. *In vitro* evidence of deaminase in kidney tissue favors the amino acids (10), but the only *in vivo* experiments yet tried have failed to show that the kidneys remove amino acids from the blood (14). Whichever is the source, the ammonia is produced from a neutral precursor.

The apparent manner in which this ammonia is manipulated by the kidney to preserve the body's neutrality may be outlined as follows, in accord with the filtration–reabsorption theory: Assuming, for purposes of illustration, that the invading acid is sulfuric acid, it will decompose body bicarbonate by the reaction,

$$H_2SO_4 + BHCO_3 = B_2SO_4 + CO_2 + H_2O$$[5]

B representing fixed base, chiefly Na and K.

The ammonia is probably extruded as ammonium bicarbonate into the glomerular filtrate as the latter passes down the tubule. The filtrate to which the ammonium bicarbonate is added already contains all the diffusible ions of the plasma, including those of the invading sulfuric acid. Consequently, with others, the ions B^+, NH_4^+, SO_4^{--}, and HCO_3^- are present in the tubular lumen. Some of the tubular cells have, as we have seen, the ability to absorb alkali bicarbonate. These selective cells absorb HCO_3^-, together with equivalent amounts of alkali cations, and let the NH_4^+ and SO_4^{--} pass into the urine. Probably reabsorption of $BHCO_3$ occurs in the distal limb of the tubule, as Richards (32) finds that the filtrate turns acid in this limb of the amphibian nephron.

[5] More logically, perhaps, the equations could be written in ionic form as $H^+ + HCO_3^- \rightleftharpoons H_2CO_3$, etc. We have, however, retained the somewhat archaic forms in order to hold in sight the essential rôle of the alkali cations in making possible the existence of the buffer anions. If the cations normally balancing HCO_3^- and (Protein)$^-$ anions were lost from the body, restoration of the buffer effects of the bicarbonate and proteinate would become impossible.

The process may be represented as follows:

B_2SO_4	+	NH_4HCO_3	=	$2\,BHCO_3$	+	$(NH_4)_2SO_4$
(Filtered from blood plasma into glomerular filtrate)		(Formed in kidney and secreted into filtrate in tubular lumen)		(Reabsorbed from tubular lumen and retained in body)		(Excreted in urine)

The net result is to replace B_2SO_4 in the body with $BHCO_3$. The regenerated $BHCO_3$ not only acts to restore the bicarbonate reserve of the body; it also acts, by shifting the equilibria of such reactions as

$$BHCO_3 + H\ (\text{Protein}) \rightarrow B\ (\text{Protein}) + H_2CO_3$$

to restore alkali to other buffers. This restoration continues until all the buffers of the body, and with them the pH, are restored to the normal state which they enjoyed before the invasion by sulfuric acid.

When the kidney is severely damaged in nephritis the abilities both to regulate bicarbonate excretion and to form ammonia decrease, and acidosis may result (9, 44).

BOUNDARIES OF PRESENT KNOWLEDGE OF RENAL FUNCTION

We have seen that renal maintenance of the volume, composition, and neutral reaction of the body fluids is explainable by developments of the filtration–reabsorption theory of Cushny and Richards, according to which the filtered substances are divided into three classes: (*1*) those which, like glucose, are completely or almost completely reabsorbed from the tubules; (*2*) those which, like water and the electrolytes, are reabsorbed as completely as may be necessary to maintain normal amounts and concentrations in the body; and (*3*) the waste products, such as urea, creatinine, and uric acid, which are not at all actively reabsorbed, although some of them slip back into the blood to varying extents by passive diffusion. The relative efficiency with which each of the substances in this third class is excreted is indicated by the volume of blood cleared of substance per minute by the kidneys.

In addition to their selective absorption from the glomerular filtrate, we have seen that the tubular cells probably elaborate the urinary ammonia, and excrete it into their lumina in exchange for alkali that they reabsorb from the glomerular filtrate. Ammonia formation and reabsorption of alkali bicarbonate appear to be the two chief mechanisms by which the kidneys maintain the neutrality of the body fluids.

Under conditions of stress the tubules furthermore appear able to assist

by direct excretion in ridding the blood of certain injected dyes and other foreign substances.

The work of Shannon has shown that in active tubular reabsorption (as of glucose) and tubular excretion (as of phenol red) the process of transfer of substance through the tubular cells appears to be accomplished by successive combination and dissociation. The mechanism in its kinetics appears to be strikingly similar to that noted for a number of hydrolytic enzymes.

One may say that the action of the glomerulus as a filter is established beyond a reasonable doubt, and that the factors controlling the speed and completeness of the filtration seem fairly well evaluated. The problem of the immediate future seems to be that of attaining some similar knowledge of the physicochemical mechanisms involved in the selective action of the tubules. This is an extraordinary function, influenced by more factors than can be estimated at present. Among these are the hormones. For example, when the pituitary fails to secrete a certain hormone or hormones, the kidneys fail to retain water to a normal degree, and diabetes insipidus results, with excretion of water up to 30 or 40 liters per day. And when the cortin of the adrenals fails, the kidneys let the sodium concentration of the plasma fall and the potassium concentration rise, until these changes alone are sufficient to threaten life. One might feel justified in abandoning explanation of tubular function to the vital forces that can never be explained. Yet advances in tubular physiology, made in the laboratories of Richards, Marshall, Smith, Shannon, and others, let one hope that the future may further clarify the mechanism that governs the volume and composition of the inner sea in which we live.

REFERENCES

(1) Austin, J. H., Stillman, Edgar, and Van Slyke, D. D.: J. Biol. Chem. **46**, 91 (1921).
(2) Chassis, H., Ranges, H. A., Goldring, W., and Smith, H. W.: J. Clin. Investigation **17**, 683 (1938).
(3) Chesley, L. C.: J. Clin. Investigation **17**, 119 (1938).
(4) Chesley, L. C., and Chesley, E. C.: Am. J. Physiol. **127**, 731 (1939).
(5) Cushny, A. R.: *The Secretion of Urine.* Longmans, Green and Co., London (1926).
(6) Dominguez, R.: Am. J. Physiol. **112**, 529 (1935).
(7) Gersh, I., and Stieglitz, E. J.: Anat. Record **58**, 349 (1933–34).
(8) Grafflin, A. L.: J. Morphol. **61**, 165 (1937).
(9) Henderson, L. J., and Palmer, W. W.: J. Biol. Chem. **17**, 305 (1914).
(10) Krebs, H. S.: Ann. Rev. Biochem. **5**, 247 (1936).
(11) Landis, E. M.: The Harvey Lectures **32**, 70 (1936–37).
(12) Marshall, E. K.: Am. J. Physiol. **99**, 77 (1931).
(13) Marshall, E. K.: Physiol. Rev. **14**, 133 (1934).
(14) Mason, Morton F.: Proc. Soc. Exptl. Biol. Med. **37**, 111 (1937).

(15) McIntosh, J. F., Møller, E., and Van Slyke, D. D.: J. Clin. Investigation **6**, 467 (1928).
(16) McLean, F. C.: J. Exptl. Med. **22**, 212, 367 (1915).
(17) McLean, F. C.: J. Exptl. Med. **26**, 181 (1917).
(18) Miller, B. F., and Dubos, R.: J. Biol. Chem. **121**, 457 (1937).
(19) Miller, B. F., and Winkler, A.: J. Clin. Investigation **15**, 489 (1936).
(20) Miller, B. F., and Winkler, A. W.: J. Clin. Investigation **17**, 31 (1938).
(21) Møller, E., McIntosh, J. F., and Van Slyke, D. D.: J. Clin. Investigation **6**, 427 (1928).
(22) Møller, E., McIntosh, J. F., and Van Slyke, D. D.: J. Clin. Investigation **6**, 485 (1928).
(23) Nash, T. P., and Benedict, S. R.: J. Biol. Chem. **16**, 117 (1922).
(24) Peters, J. P.: *Body Water.* Charles C. Thomas, Springfield, Illinois (1935).
(25) Rehberg, P. B.: Biochem. J. **20**, 447 (1926).
(26) Rhoads, C. P.: Am. J. Physiol. **109**, 324 (1934).
(27) Richards, A. N.: The Harvey Lectures **30**, 93 (1934–35).
(28) Richards, A. N.: Bull. N. Y. Acad. Med. **14**, 5 (1938).
(29) Richards, A. N.: "Methods and Results of Direct Investigations of the Function of the Kidney," Beaumont Lecture, Baltimore, 1929.
(30) Richards, A. N.: Proc. Roy. Soc. (London) **B126**, 398 (1938).
(31) Richards, A. N., Bordley, J., and Walker, A. M.: J. Biol. Chem. **101**, 179 (1933).
(32) Richards, A. N., Walker, A. M., Hudson, C. L., Findley, Thos., Jr., Montgomery, H., and Pierce, J. A.: Am. J. Physiol. **118**, 111–73 (1937).
(33) Sendroy, J., Jr., Seelig, S., and Van Slyke, D. D.: J. Biol. Chem. **106**, 463, 479 (1934).
(34) Shannon, James: Physiol. Rev. **19**, 63 (1939).
(35) Shannon, J. A., and Fisher, S.: Am. J. Physiol. **122**, 765 (1938).
(36) Smith, H. W.: *The Physiology of the Kidney.* Oxford University Press, New York (1937).
(37) Smith, H. W.: "Studies in the Physiology of the Kidney," Porter Lectures, Lawrence, Kansas, 1939.
(38) Smith, H. W.: Ann. Rev. Physiol. **1**, 503 (1939).
(39) Van Slyke, D. D.: J. Biol. Chem. **73**, 695 (1927).
(40) Van Slyke, D. D., and Cullen, G. E.: J. Biol. Chem. **19**, 141 (1914).
(41) Van Slyke, D. D., and Hawkins, J. A.: J. Biol. Chem. **83**, 51 (1929).
(42) Van Slyke, D. D., Hiller, Alma, and Miller, B. F.: Am. J. Physiol. **113**, 611 (1935).
(43) Van Slyke, D. D., Hiller, Alma, and Miller, B. F.: Am. J. Physiol. **113**, 629 (1935).
(44) Van Slyke, D. D., Linder, G. C., Hiller, Alma, Leiter, L., and McIntosh, J. F.: J. Clin. Investigation **2**, 255 (1926).
(45) Van Slyke, D. D., Rhoads, C. P., Hiller, Alma, and Alving, A. S.: Am. J. Physiol. **109**, 336 (1934).
(46) Van Slyke, D. D., Rhoads, C. P., Hiller, Alma, and Alving, A.: Am. J. Physiol. **110**, 387 (1934).
(47) Van Slyke, D. D., Stillman, E., Møller, E., Ehrlich, E., McIntosh, J. F., Leiter, L., MacKay, E. N., Hannon, R. R., Moore, N. S., and Johnston, C.: Medicine **9**, 257 (1930).
(48) Van Slyke, D. D., and Zacharias, G.: J. Biol. Chem. **19**, 181 (1914).

PHOTOREACTIONS SENSITIZED BY THE HALOGENS[1]

J. W. T. SPINKS

Department of Chemistry, University of Saskatchewan, Saskatoon, Canada

Received November 11, 1938

In a series of photochemical reactions, the halogens are the active light-absorbing component. The halogens enter directly into some of the reactions, in others they merely act as sensitizers. As is well known, sensitized reactions may take place in the gaseous, liquid, or solid phase. This paper will deal mainly with homogeneous gas reactions sensitized by the halogens.

When a molecule absorbs light, one of a number of things may happen; the molecule may be excited, dissociated, or predissociated, or possibly even ionized. Fortunately the nature of the absorption spectrum often indicates what happens to the molecule when it absorbs a light quantum, that is, the absorption spectrum indicates the nature of the primary photochemical act.

According to spectroscopic evidence, the primary photochemical process in chlorine is dissociation. In bromine above 5107 Å. and in iodine above 4989 Å. the primary process is excitation, but evidence points to this excitation being converted into dissociation by collision if foreign gases are present, which is always the case in photosensitization experiments (7, 40). Therefore, the photochemical reactions sensitized by the halogens are probably always initiated by free atoms. Examples discussed in this paper show that this assumption provides a satisfactory explanation of the facts.

THE ACTION OF CHLORINE

1. The decomposition of ozone

Chlorine sensitizes the decomposition of ozone, both thermally and photochemically, oxygen being produced. With less than 7 per cent ozone, the quantum yield is about 3. The reaction is zero order with respect to ozone concentration until the ozone is practically all decomposed, when the rate suddenly increases. A similar sudden increase in rate is

[1] This paper was read at the Meeting of the American Association for the Advancement of Science which was held in Ottawa, Canada, June, 1938.

observed for nitrogen trichloride (1, 4, 24). When higher ozone concentrations are used, a chain reaction appears and quantum yields up to 60 have been observed. With increasing ozone concentration or decreasing chlorine concentration, the reaction approaches the $I_{\text{abs.}}^{1/2}$ type. At the same time, the reaction becomes sensitive to small temperature changes and to changes in surface (16). The formation of ClO_2 and ClO_3 as intermediate products can be shown, and under suitable conditions Cl_2O_6 separates out on the walls of the vessel as a red liquid (24, 27). Chlorine heptoxide is also formed under certain conditions. Chlorine heptoxide is relatively inert, but ClO_3 undergoes various reactions, such as thermal and photosensitized decomposition reactions, with ozone (27). As a simplified mechanism for this reaction, we suggest:

$$Cl_2 + h\nu \rightarrow Cl + Cl \tag{1}$$

$$Cl + O_3 \rightarrow ClO + O_2 \tag{2}$$

$$ClO + O_3 \rightarrow ClO_2 + O_2 \tag{3}$$

$$\left.\begin{aligned} ClO_2 + O_3 &\rightarrow ClO_3 + O_2 \quad (4) \\ ClO_3 + O_3 &\rightarrow ClO_2 + 2O_2 \quad (5) \end{aligned}\right\}\text{chain}$$

Reactions 4 and 5 are known definitely to take place.

$$2ClO_3 \rightarrow Cl_2O_6 \tag{6}$$

Goodeve and Todd (14) state that this oxide exists mainly as ClO_3 in the gas phase but as Cl_2O_6 in the liquid phase. Instead of reaction 2, Rollefson and Burton (26) propose

$$Cl + O_3 \rightarrow ClO_3$$

They also suggest that ClO_4 is formed, to account for the formation of chlorine heptoxide.

$$ClO_3 + O_3 \rightarrow ClO_4 + O_2$$

$$ClO_4 + ClO_3 \rightarrow Cl_2O_7$$

It will be realized that when the various possible thermal and photo reactions of the intermediates are taken into account, the reaction is scarcely as simple as Bonhoeffer supposed (quantum yield = 2) in 1923.

2. The decomposition of chlorine monoxide

Chlorine photosensitizes the decomposition of chlorine monoxide (2), the rate of decomposition being proportional to the amount of light absorbed. The quantum yield is 2, the same as for the direct photodecomposition. This can be very satisfactorily explained by the following

mechanism: From spectroscopic evidence the mechanism of the direct decomposition is

$$Cl_2O + h\nu \rightarrow ClO + Cl$$

$$Cl + Cl_2O \rightarrow Cl_2 + ClO$$

followed by reactions of ClO such as

$$2ClO \rightarrow Cl_2 + O_2$$

The mechanism suggested for the sensitized reaction is

$$Cl_2 + h\nu \rightarrow Cl + Cl$$

$$Cl + Cl_2O \rightarrow Cl_2 + ClO$$

$$2ClO \rightarrow Cl_2 + O_2$$

leading to the same quantum yield as before.

3. The decomposition of ClO_2 *and* ClO_3

Experiments have not yet been done on the chlorine-sensitized decomposition of ClO_2, although there is no doubt that it would go as readily as the bromine-sensitized decomposition to be described in the next section. Preliminary experiments have recently been done on the sensitized decomposition of ClO_3 (18). Gaseous ClO_3 absorbs in the ultraviolet with a threshold at about 3500 Å. (13). Consequently, with the maximum pressure of ClO_3 possible at 3°C., very little light is absorbed at 3650 Å. On adding chlorine and insolating with 3650 Å., a sensitized reaction takes place, the quantum yield being about 0.7. In calculating this quantum yield, the overall photoreaction was assumed to be

$$2ClO_3 \rightarrow Cl_2 + 3O_2$$

and the reaction was followed by the pressure change on a glass spoon gauge. More work has to be done on this reaction before a mechanism can be put forward.

4. The decomposition of nitrogen trichloride

The chlorine-sensitized decomposition of nitrogen trichloride has been extensively studied (15). The reaction is of zero order with respect to nitrogen trichloride until the reaction is nearly complete; then a semi-explosive reaction occurs. With a high chlorine pressure, the quantum yield approaches a limiting value of 2. The primary action is

$$Cl_2 + h\nu \rightarrow Cl + Cl \tag{1}$$

which is followed by

$$NCl_3 + Cl \rightarrow NCl_2 + Cl_2 \tag{2}$$

The NCl_2 then reacts, producing chains that are short at ordinary temperature.

$$NCl_2 + NCl_3 \rightarrow N_2 + 2Cl_2 + Cl \quad (3)$$

While the quantum yield is independent of 2.5-fold variation of surface, there are some indications that the state of the surface may enter indirectly as a factor in the decomposition. In vessels that have been freshly washed, quantum yields 50 per cent higher than usual are occasionally observed. Chain rupture is supposed to be due to

$$Cl + NCl_3 + X \rightarrow NCl_4 + X^* \quad (4)$$

$$2NCl_4 + \text{(surface catalyst)} \rightarrow N_2 + 4Cl_2 \quad (5)$$

It can be shown that, using this mechanism,

$$(NCl_2)\ (NCl_3) = \text{constant}$$

The semi-explosive termination might therefore be due to the building up of the intermediate product, NCl_2, as (NCl_3) approaches zero. The sudden increase in (NCl_2) might cause other heat-producing reactions such as

$$NCl_2 + NCl_2 \rightarrow N_2 + 2Cl_2$$

to become important, producing a sudden rise in pressure.

5. *Oxidations*

Chlorine also acts as sensitizer in a number of extremely interesting oxidation reactions of simple carbon compounds (29). For example, chlorine and carbon monoxide unite to form phosgene, but in the presence of oxygen the formation of phosgene is practically completely suppressed and a sensitized formation of carbon dioxide takes place (for a summary see, for example, reference 5). It seems probable that the radical COCl is formed as an intermediate. This radical reacts preferentially with oxygen, forming a peroxide which gives rise to a chain reaction in which the halogen is reformed and the organic substance oxidized with a quantum yield that may amount to about 10^3 molecules per quantum.

$$Cl_2 + h\nu \rightarrow 2Cl \quad (1)$$

$$Cl + CO + M \rightarrow COCl + M \quad (2)$$

$$COCl + O_2 \rightarrow CO_3Cl \quad (3)$$

$$CO_3Cl \rightarrow CO_2 + ClO \quad (4)$$

$$ClO + CO \rightarrow CO_2 + Cl \quad (5)$$

The chain carrier in this case is probably ClO.

Chloroform and chlorine give carbon tetrachloride in the light, but in the presence of oxygen phosgene is formed almost quantitatively with a quantum yield of about 100 (9, 31).

$$Cl_2 + h\nu \rightarrow Cl + Cl \tag{1}$$

$$Cl + CHCl_3 \rightarrow CCl_3 + HCl \tag{2}$$

$$CCl_3 + O_2 \rightarrow CO_2Cl_3 \tag{3}$$

$$CO_2Cl_3 \rightarrow COCl_2 + ClO \tag{4}$$

$$ClO + CHCl_3 \rightarrow HCl + COCl_2 + Cl \tag{5}$$

Here the radical CCl_3 appears to form a peroxide, which yields $COCl_2$ and a chain carrier ClO.

Experiments recently carried out in this laboratory indicate that reaction 5 actually does take place when chloroform reacts with a ClO radical, prepared by shining light on ClO_2.

$$ClO_2 + h\nu \rightarrow ClO + O$$

This and similar reactions are being investigated in detail.

Franke and Schumacher (12) have investigated the kinetics of the chlorine-sensitized oxidation of trichlorobromomethane. Phosgene is produced, and quantum yields up to 400 have been measured.

$$Cl_2 + h\nu \rightarrow Cl + Cl \tag{1}$$

$$Cl + CCl_3Br \rightarrow CCl_3 + BrCl \tag{2}$$

$$CCl_3 + O_2 \rightarrow CO_2Cl_3 \tag{3}$$

$$CO_2Cl_3 \rightarrow COCl_2 + ClO \tag{4}$$

$$ClO + CCl_3Br \rightarrow COCl_2 + \begin{cases} Cl_2 + Br \\ BrCl + Cl \end{cases} \tag{5}$$

$$CO_2Cl_3 + ClO \rightarrow COCl_2 + O_2 + Cl_2 \tag{6}$$

A mixture of tetrachloroethylene and chlorine forms hexachloroethane on illumination, but in the presence of oxygen trichloroacetyl chloride is formed (11).

$$Cl_2 + h\nu \rightarrow Cl + Cl \tag{1}$$

$$Cl + C_2Cl_4 \rightarrow C_2Cl_5 \tag{2}$$

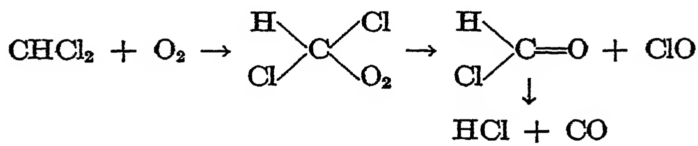
$$C_2Cl_5 + O_2 \rightarrow C_2Cl_5O_2 \tag{3}$$

$$C_2Cl_5O_2 \rightarrow CCl_3 \cdot COCl + ClO \tag{4}$$

$$ClO + C_2Cl_4 \rightarrow CCl_3 \cdot COCl + Cl \tag{5}$$

Trichloroethylene similarly gives dichloroacetyl chloride (22).

In order to clarify some of the foregoing oxidations, Brenschede and Schumacher investigated the reactions of methane, methyl chloride, methylene chloride, and formaldehyde with oxygen and chlorine and light (6). Methane yields mainly carbon dioxide; methyl chloride yields more carbon monoxide; methylene chloride yields carbon monoxide and phosgene in the ratio 10:1. In the last reaction the radical $CHCl_2$ apparently reacts with oxygen, forming a peroxide which yields formyl chloride and a chain carrier. Chlorine is practically not used up in the reaction.



Formyl chloride undergoes either a thermal or chlorine-sensitized decomposition, yielding hydrogen chloride and carbon monoxide (28). The formation of a peroxy compound in the methane reaction was shown by streaming experiments in which the products were frozen out in a trap at −95°C. A small amount of oily liquid was obtained which gave a positive peroxy test with titanic acid.

In each of these oxidation reactions, an intermediate compound with a trivalent carbon is formed, either by abstracting an atom such as hydrogen or bromine from an organic compound or by adding on a halogen atom to a C=C double bond. The radical formed reacts with oxygen, producing a peroxide which gives rise to a chain reaction in which the halogen is reformed and various oxidation products result. It is also typical that the sensitized oxidations only take place at a temperature below 150°C., owing to the ease of decomposition of the peroxide. The following equations summarize the overall reactions of the radicals:

$$COCl + O_2 = CO_2 + ClO$$

$$CCl_3 + O_2 = COCl_2 + ClO$$

$$C_2Cl_5 + O_2 = CCl_3COCl + ClO$$

$$C_2HCl_4 + O_2 = CHCl_2COCl + ClO$$

$$CHCl_2 + O_2 = CO + HCl + ClO$$

A general reaction mechanism can be written for the above oxidations:

$$Cl_2 + h\nu \rightarrow Cl + Cl \tag{1}$$

$$Cl + X \rightarrow R + HCl, \text{ or} \tag{2}$$

$$Cl + X^1 \rightarrow R^1$$

$$R \text{ (or } R^1) + O_2 \rightarrow P \tag{3}$$

$$P \rightarrow \text{Products} + ClO \tag{4}$$

$$ClO + X \rightarrow \text{Products} + Cl \tag{5}$$

X is a halogenated hydrocarbon (X^1 unsaturated); R (or R^1) is the radical formed; and P is a peroxy compound.

THE ACTION OF BROMINE

1. The decomposition of ozone

The thermal decomposition of ozone is catalyzed by bromine, the oxide $(Br_3O_8)_n$ being formed as an intermediate compound (20). Other oxides, Br_2O and BrO_2, are known and their formation may precede the formation of $(Br_3O_8)_n$. Using low concentrations of bromine and ozone, a photoreaction can be measured, the reaction being a chain reaction with a quantum yield of about 30 (4, 32). It is interesting to note that the quantum yields at 3650 and 5460 Å. are the same, indicating that practically every bromine molecule absorbing a quantum at 5460 Å. dissociates into atoms. The chain is probably propagated by an unstable oxide of bromine, possibly BrO. Indeed, the decompositions of ozone photosensitized by chlorine and bromine, respectively, resemble one another in many respects.

2. The decomposition of chlorine monoxide

Chlorine monoxide undergoes a bromine-sensitized decomposition using light of wave length 5460 Å. (7). The reaction involves short chains and some chlorine dioxide is formed. The quantum yield is 4.3 at 19°C., and the ratio of the quantum yields in the direct and sensitized reactions is practically unity. This suggests that, apart from the primary act of light absorption, the sensitized and unsensitized reactions have essentially similar mechanisms. In all probability, 5460 Å. produces bromine atoms, either directly or indirectly, and a possible mechanism is therefore;

$$Br_2 + h\nu \rightarrow (Br_2^*) \rightarrow 2Br \tag{1}$$

$$Br + Cl_2O \rightarrow BrCl + ClO \tag{2}$$

$$2BrCl = Br_2 + Cl_2 \tag{3}$$

Quantum yields greater than 2 and the formation of ClO_2 might be due to

$$ClO + Cl_2O \rightarrow Cl + Cl_2 + O_2 \tag{4}$$

and

$$ClO + Cl_2O \rightarrow ClO_2 + Cl_2 \tag{5}$$

Chains could be terminated by

$$Cl + Cl + M \rightarrow Cl_2 + M$$

$$ClO + ClO \rightarrow Cl_2 + O_2$$

$$Cl + ClO + M \rightarrow Cl_2O + M$$

$$Cl + ClO_2 \rightarrow Cl_2 + O_2$$

The relative importance of these reactions has not yet been determined. It is, of course, possible that ClO may react with bromine,

$$ClO + Br_2 \rightarrow BrO + BrCl$$

but, thus far, experiments to prove or disprove this suggestion are lacking.

3. *The decomposition of chlorine dioxide*

Bromine also sensitizes the decomposition of chlorine dioxide (33). At 15°C. the reaction is accompanied by a pressure decrease and the formation of Cl_2O_6. If insolation is continued, the pressure subsequently increases, owing to the photosensitized decomposition of Cl_2O_6 (or ClO_3). At 30°C. the reaction is accompanied by an increase in pressure and the formation of chlorine and oxygen. By introducing water into the reaction cell, troublesome side reactions can be eliminated and the effect of varying different factors studied. It is found that the quantum yields at 3650 and 5460 Å. are equal, indicating that the reaction at 5460 Å. probably also proceeds by way of bromine atoms and not by way of an excited molecule. The quantum yields for the sensitized and unsensitized reactions are also equal, indicating that, apart from the primary act of light absorption, the mechanisms for the two reactions are similar. In the direct reaction the mechanism probably is:

$$ClO_2 + h\nu \rightarrow ClO + O$$

$$ClO_2 + O + M \rightarrow ClO_3 + M$$

followed by reactions of ClO and ClO_3.

For the sensitized reaction Schumacher (30) has suggested a mechanism involving bromine atoms:

$$Br_2 + h\nu \rightarrow 2Br$$

$$Br + ClO_2 \rightarrow BrCl + O_2$$

$$2BrCl = Br_2 + Cl_2$$

However, since the main product of the reaction is Cl_2O_6, this mechanism has to be rejected. Another possibility would be

$$Br_2 + h\nu \rightarrow 2Br$$

$$Br + ClO_2 \rightarrow BrO + ClO$$

$$BrO + ClO_2 \rightarrow ClO_3 + Br$$

which would give the same products as the direct photoreaction. In view of the existence of other oxides of bromine, it does not seem to be particularly objectionable to postulate the existence of BrO.

4. The decomposition of ClO_3

The decomposition of ClO_3 is sensitized by bromine, but thus far quantitative experiments are lacking (33).

5. The decomposition of dibromotetrachloroethane

The bromine-sensitized decomposition of gaseous dibromotetrachloroethane has been studied (8). Using 4360 Å. and a temperature of about 150°C. tetrachloroethylene is produced, the quantum yield varying from 1 to 40. The mechanism is essentially similar to that of the iodine-sensitized decomposition of ethylene iodide in carbon tetrachloride solution.

$$Br_2 + h\nu \rightarrow 2Br$$

$$Br + C_2Cl_4Br_2 \rightarrow C_2Cl_4Br + Br_2$$

$$C_2Cl_4Br \rightleftarrows C_2Cl_4 + Br$$

$$Br \xrightarrow{\text{wall}} \tfrac{1}{2}Br_2$$

$$Br + Br + M \rightarrow Br_2 + M$$

6. The decomposition of dibromodichloroethane

Bromine sensitizes the decomposition of dibromodichloroethane (23), the quantum yield being about 12.

7. The decomposition of bromophosgene

The decomposition of bromophosgene is accelerated by bromine plus light, but the quantum yield is small (4).

8. The decomposition of nitrous oxide

Bromine also photosensitizes the decomposition of nitrous oxide at 650°C. (17).

$$Br_2 + h\nu \rightarrow Br + Br$$

$$N_2O + Br \rightarrow BrO + N_2$$

$$N_2O + BrO \rightarrow Br + O_2 + N_2$$

9. *Oxidations*

Bromine sensitizes the oxidation of bromotrichloromethane to phosgene (12). Using 4360 Å. and (CCl_3Br) < 4(Br_2), the quantum yield does not exceed 4, and thereby differs quite markedly from the yield for corresponding chlorine-sensitized reaction. The first steps in the reaction are:

$$Br_2 + h\nu \rightarrow Br + Br \tag{1}$$

$$CCl_3Br + Br \rightarrow CCl_3 + Br_2 \tag{2}$$

$$CCl_3 + O_2 \rightarrow COCl_2 + ClO \tag{3}$$

The lower quantum yield in this case is thought to be due to the removal of the chain carrier by the reaction,

$$ClO + Br_2 \rightarrow BrO + BrCl \tag{4}$$

It is then assumed that BrO reacts with bromotrichloromethane to form phosgene, without regenerating the chain carrier.

$$CCl_3Br + BrO \rightarrow COCl_3 + Br_2 \tag{5}$$

$$2COCl_3 \rightarrow 2COCl_2 + Cl_2 \tag{5a}$$

This mechanism leads to a limiting quantum yield of 4. If (CCl_3Br) > 4(Br_2), ClO reacts with CCl_3Br, leading to chain formation and a quantum yield greater than 4 (see equation 5 in the chlorine-sensitized reaction).

In contrast to the corresponding chlorine-sensitized reaction, bromine does not sensitize the oxidation of carbon monoxide (21). A number of further oxidations sensitized by bromine have been carried out in carbon tetrachloride solution, but will not be dealt with here (10, 19).

THE ACTION OF IODINE

While iodine sensitizes many reactions, very few of these have been investigated in the gas phase. The only one that will be mentioned is the decomposition of nitrous oxide, which is accelerated by light in the presence of iodine (17).

From the above survey, it is clear that all the halogen-photosensitized gaseous reactions thus far known can be explained by chemical mechanisms involving halogen atoms.

SUMMARY

Spectroscopic evidence indicates that the sensitizing action of chlorine is practically always due to chlorine atoms. For bromine, when λ is less than 5107 Å., reaction will be by way of bromine atoms. Experiments using 5107–6290 Å. also probably involve bromine atoms, although the possibility of reaction due to excited bromine molecules cannot be wholly ruled out. Iodine atoms are produced in photosensitization experiments

involving iodine and λ less than 4989 Å. Experimental evidence also favors the production of iodine atoms using λ up to 6200 Å.

Mechanisms in accordance with the above are advanced for a number of chlorine-sensitized decomposition reactions and for a number of chlorine-sensitized oxidation reactions. In the oxidation reactions, an intermediate compound with a trivalent carbon is formed which reacts with oxygen forming a peroxide, giving rise to a chain reaction. Reactions sensitized by bromine and iodine are also discussed.

It appears that all the halogen-sensitized photoreactions thus far known can be explained by chemical mechanisms involving halogen atoms.

REFERENCES

(1) Allmand and Spinks: J. Chem. Soc. **1931**, 1652; **1932**, 599.
(2) Bodenstein and Kistiakowsky: Z. physik. Chem. **116**, 371 (1925).
(3) Bonhoeffer: *Grundlagen der Photochemie.* Steinkopff, Leipzig (1933).
(4) Bonhoeffer: Z. Physik **13**, 94 (1923).
(5) Brenschede: Z. physik. Chem. **B41**, 237 (1938).
(6) Brenschede and Schumacher: Z. physik. Chem. **A177**, 245 (1936).
(7) Brown and Spinks: Can. J. Research **B15**, 113 (1937).
(8) Carrico and Dickinson: J. Am. Chem. Soc. **57**, 1343 (1935).
(9) Chapman: J. Am. Chem. Soc. **57**, 416 (1935).
(10) Daniels: J. Am. Chem. Soc. **57**, 2240 (1935).
(11) Dickinson and Leermakers: J. Am. Chem. Soc. **54**, 3853 (1934).
(12) Franke and Schumacher: Z. physik. Chem. **B42**, 297 (1939).
(13) Goodeve and Richardson: Trans. Faraday Soc. **33**, 453 (1937).
(14) Goodeve and Todd: Nature **132**, 514 (1933).
(15) Griffiths and Norrish: Proc. Roy. Soc. (London) **A130**, 591 (1931); **A135**, 69 (1932).
(16) Heidt, Kistiakowsky, and Forbes: J. Am. Chem. Soc. **55**, 223 (1933).
(17) Hinshelwood and Musgrave: Proc. Roy. Soc. (London) **A137**, 25 (1932).
(18) Kalina and Spinks: Can. J. Research **B16**, 381 (1938).
(19) Koblitz, Meissner, and Schumacher: Ber. **B70**, 1080 (1937).
(20) Lewis and Schumacher: Z. physik. Chem. **A137**, 462 (1928); J. Am. Chem. Soc. **54**, 3185 (1932).
(21) Livingstone: J. Phys. Chem. **34**, 2121 (1930).
(22) Muller and Schumacher: Z. physik. Chem. **B37**, 365 (1937).
(23) Muller and Schumacher: Z. physik. Chem. **B42**, 327 (1939).
(24) Norrish and Neville: J. Chem. Soc. **1934**, 1864.
(25) Rollefson: *Photochemistry of the Halogens.* Paris, Hermann et Cie. (1938).
(26) Rollefson and Burton: *Photochemistry,* p. 288. Prentice-Hall, New York (1939).
(27) Rollefson and Byrns: J. Am. Chem. Soc. **56**, 364, 1250, 2245 (1934).
(28) Rollefson and Krauskopf: J. Am. Chem. Soc. **54**, 2542 (1934).
(29) Schumacher: Z. Elektrochem. **42**, 522 (1936).
(30) Schumacher: Trans. Electrochem. Soc. **71**, 409 (1937).
(31) Schumacher and Wolff: Z. physik. Chem. **B26**, 453 (1934).
(32) Spinks: Nature **128**, 548 (1931).
(33) Spinks and Porter: J. Am. Chem. Soc. **56**, 264 (1934).

SYMPOSIUM ON X-RAY STUDIES OF SUBSTANCES OF HIGH MOLECULAR WEIGHT[1]

INTRODUCTION TO THE SYMPOSIUM

MAURICE L. HUGGINS

Eastman Kodak Company, Rochester, New York

Received October 3, 1939

Twenty-six years have elapsed since W. H. Bragg and W. L. Bragg reported the first determination of the structure of a crystal with the aid of x-rays. Although the fundamental principles of the interaction of x-rays and crystals remain the same, there have been great improvements both in the technique of obtaining the data and in the technique of interpreting them. The structural details of complex organic molecules can now be successfully studied, the results checking well, where comparison is possible, with those obtained by other methods, such as electron diffraction, band spectrum, and dipole moment studies.

In all of the earlier x-ray diffraction work and in much that is done today, the x-ray data are used to study structural *regularities*. Nowadays, however, they are also being used, more and more, to give information regarding the *irregularities*—for example, random replacements of one kind of atom by another, oscillations and rotations of atomic groups and even of whole molecules, irregular distributions of alternative atomic arrangements having practically the same energy, and molecular orientation and structure in fibers and other not entirely crystalline materials. X-ray methods are not ideally suited to such studies and in many cases cannot tell us the whole story, yet they do furnish a certain amount of information which is at present obtainable in no other way. In order to use the x-ray data to the fullest extent, it is usually also necessary, with the more complex materials, to utilize information from various other fields and to reason to some extent by analogy with other comparable structures. This is, of course, quite proper and good scientific practice, provided it is done with sufficient care.

[1] This Symposium was held at the Ninety-eighth Meeting of the American Chemical Society in Boston, Massachusetts, September, 1939, under the auspices of the Divisions of Physical and Inorganic Chemistry, Colloid Chemistry, Organic Chemistry, Paint and Varnish Chemistry, and Rubber Chemistry of the American Chemical Society.

As the structures being studied become more complicated and as the necessity for including evidence from other sources increases, the conclusions reached depend more and more on the judgment of those drawing them. As a rule, non-specialists in the field can no longer judge adequately the validity of these conclusions by an inspection of the papers reporting the work. Moreover, for reasons of rigor and conciseness, the methods and frequently—indeed, almost invariably—even the results are reported in language which is unintelligible to the average reader.

Largely for these reasons, this symposium has been arranged. It brings before you some of the more important results which have been obtained with the aid of x-ray diffraction.

X-ray methods can be applied to very diverse problems. Those who work in the field often see applications to problems which, at first sight, are quite unrelated. This is especially true of industrial applications. It is hoped that valuable ideas and suggestions will be obtained from a perusal of these papers. The scope of the symposium is restricted to studies of substances of high molecular weight. These include such materials as rubber, cellulose and its derivatives, synthetic linear polymers, and proteins, in all of which x-ray evidence is of considerable technical as well as theoretical interest. Proteins are of considerable interest to biologists.

Although the papers treat primarily of substances composed of what we might call "*giant* molecules," some do, of course, discuss researches on substances composed of *small* molecules, whenever these help to furnish a suitable background of information.

Strictly speaking, most minerals—one might even say most inorganic compounds—should perhaps be included in the term "substances of high molecular weight." In this symposium, however, the discussion of inorganic substances is limited to glasses. These possess certain elements of irregularity which make the problems involved in the interpretation of the x-ray data similar in many respects to those met with in work on the other substances to be discussed.

THE INVESTIGATION OF SYNTHETIC LINEAR POLYMERS BY X-RAYS[1]

C. S. FULLER

Bell Telephone Laboratories, New York, New York

Received October 3, 1939

CONTENTS

I. INTRODUCTION

The purpose of this paper is to review the results which have been obtained to date by the x-ray and electron diffraction study of synthetic organic linear polymers[2]. The subject matter will be confined to the truly synthetic compounds, and work relating to the derivatives of the natural linear polymers will be omitted. The latter studies for the most part relate to derivatives of cellulose, and these have already been reviewed by Sisson (68). Discussion of the techniques employed will likewise be

[1] Presented at the Symposium on X-ray Studies of Substances of High Molecular Weight, held at the Ninety-eighth Meeting of the American Chemical Society in Boston, Massachusetts, September, 1939, under the auspices of the Divisions of Physical and Inorganic Chemistry, Colloid Chemistry, Organic Chemistry, Paint and Varnish Chemistry, and Rubber Chemistry of the American Chemical Society.

[2] This report is written more from the x-ray standpoint. In general, x-rays are better adapted to the study of thick specimens, whereas electrons are more useful for the examination of thin films and surfaces.

omitted, since it will be necessary to conserve space in order to give a unified picture of the results on synthetic polymers in the allotted space. Readers interested in this part of the subject are referred to the literature and to standard works on the subject (1, 15, 30, 48, 42). In order to secure a better appreciation of the meaning of the investigations that have been made, it will be helpful to consider certain related topics before undertaking a study of the results themselves. We shall consider first the various ways in which x-rays can be employed in the examination of high polymeric substances.

II. CLASSIFICATION OF X-RAY METHODS

As in the case of the low molecular compounds there are a number of ways in which x-rays can be applied to the study of high polymeric substances. The various methods may be divided conveniently into three categories, depending on the object of the investigation. It may be desired (*1*) to identify a given substance, (*2*) to study the effect of certain variables on it, or, finally, (*3*) to unravel its inner molecular structure. In the first case, the procedure is the usual one of comparing the diffraction pattern of the unknown with patterns furnished by compounds of known constitution. Although this use of x-rays is of particular value in the case of high molecular substances because of the simple technique involved and because conventional methods are often difficult to apply to these substances, this use of x-rays will not concern us here.

The second category mentioned above is that in which the effect of a given variable on the system is the object of study. According to this method the investigator studies the effect of temperature, pressure, or a second substance on the material in question. The effect of elongation on the x-ray pattern of rubber or the effect of sodium hydroxide on the x-ray pattern of cellulose are typical examples. Such studies are valuable in explaining the influence of external agents on the inner molecular or crystalline nature of the substance under consideration, and much of the work on the natural polymers falls in this classification.

In most of the studies reported here we shall be concerned with the last of the methods mentioned above, that is, with work designed to give information on the structural features of the substances under study. The specific utility of x-rays in this direction has been amply demonstrated by innumerable investigations on inorganic and organic compounds of low molecular weight. In applying this method to substances of high molecular weight the same rules and precautions hold as with these low molecular compounds, but it is perhaps even more essential for the investigator to be aware of its limitations and pitfalls.

In investigations with low molecular compounds single crystals are

generally available, but in the case of the high polymers we are limited to the investigation of polycrystalline, mesomorphic, or amorphous masses. Accordingly, the results are less definite and more difficult of interpretation than in the case of the simpler substances. For this reason it is necessary to consider chemical evidence along with x-ray data. In structural studies on high polymers in the past one major difficulty has been the lack of definite knowledge regarding the chemical structure of these compounds. In the last decade considerable progress has been made in the synthesis of polymeric substances of known structure. There can be little doubt that, aided by these new data, the x-ray method will continue to contribute important information regarding the constitution of high polymeric substances.

In addition to the purely structural studies, x-rays also furnish considerable information on the colloidal nature of high polymers. Thus suitable techniques allow estimates to be made of the crystal particle size, the proportion of crystalline to amorphous material present, and finally the orientation of the crystalline components relative to the surfaces of the specimen. The importance of these results in correlating the physical properties of these systems with their colloidal structure will be considered later.

III. LOW MOLECULAR CHAIN COMPOUNDS

No consideration of the x-ray work on linear high polymers can neglect the extensive results which have been accumulated on the lower molecular chain compounds. Without going into detail, some of the important conclusions from these studies may now be reviewed. The first significant result is the proof that the long aliphatic chains in these compounds assume an extended tetrahedral zigzag form. Such a configuration (figure 4a) had been suggested in 1917 by Langmuir (40) from surface studies. Later Piper and Grindley (59), Müller (50), Shearer (66), and Müller and Shearer (52), working with soaps and fatty acids, found that a parallel arrangement of chains in which the atoms were arranged in zigzag fashion accounted well for the x-ray results on these compounds. Subsequent work by Müller (51), Hengstenberg (28), and others has established beyond all doubt the essential correctness of this configuration of the hydrocarbon chains in the crystalline aliphatic compounds.

A second significant conclusion from the work on the chain compounds of lower molecular weight is that in all of these compounds the chains are packed together in the crystals as parallel rods of approximately the same effective cross section (18.5 Å.^{2}). These chains may be perpendicular or at a definite angle to the planes containing the ends of the molecules. Figure 1 illustrates the angles which have been reported for various paraffins,

aliphatic alcohols, acids, and esters. The same substance under the proper conditions may show several different angles. It is on the basis of this difference in angle or "tilt" of the molecular chains that various polymorphic forms of certain of these compounds have been differentiated[3].

Recently Kohlhaas and Soremba (33) have investigated carefully prepared single crystals of *n*-triacontane ($C_{30}H_{62}$) and Kohlhaas (32) has examined a single crystal of cetyl palmitate; they have thus verified again the main conclusions in regard to the structures of these compounds. In an important paper T. Schoon (65) has presented a general theory of crystal formation of these aliphatic chain compounds. A brief description of his conception will be of interest, since it has a definite bearing on the results on synthetic polymers to be considered later. Schoon points out that

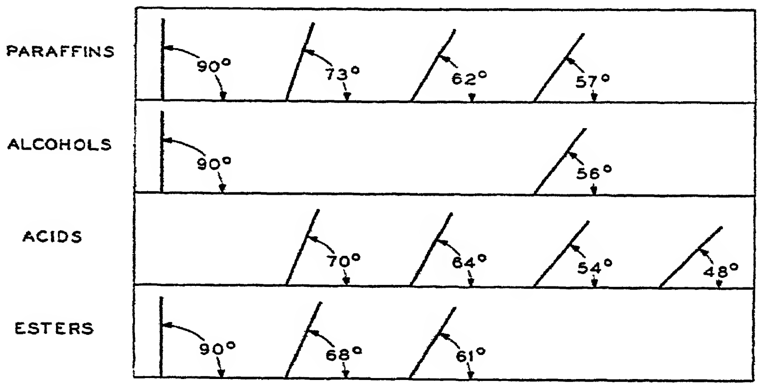


Fig. 1. Inclination of hydrocarbon chains in various aliphatic compounds

simple geometrical relationships exist between lattices of the various polymorphic crystal forms found in the case of most of these chain substances. The various forms are pictured as originating through a molecular gliding process, as illustrated in figure 2. Thus, an orthorhombic form of crystal lattice such as that pictured in figure 2a may transform into other stable crystalline forms simply by a uniform gliding of rows of chains, as shown in figure 2, b, c, and d. A type of glide such as in b or c gives rise to a monoclinic cell, whereas the type shown in figure 2d results in a triclinic cell[4]. In this manner Schoon has been able to correlate a large

[3] It should be pointed out that two types of modifications are distinguished in long-chain aliphatic compounds;—the rotating (5) and the non-rotating. The rotating forms always show the perpendicular chain arrangement; the non-rotating forms may be perpendicular or tilted.

[4] So far this type has not been observed.

amount of x-ray data on these compounds in a fairly quantitative way. For example, stearic acid has been found in a form having a unit cell a = 5.62 Å., b = 7.54 Å., and β = 61.2° and in a form having a = 9.46 Å., b = 4.96 Å., and β = 54.2° (65). Schoon's conception accounts for the first form as a glide of successive (100) planes a distance of two chain atoms, as shown in figure 2b. The second form corresponds to a glide of successive (010) planes a distance of four chain atoms, as shown in figure 2c. Further reference to Schoon's theory will be made below.

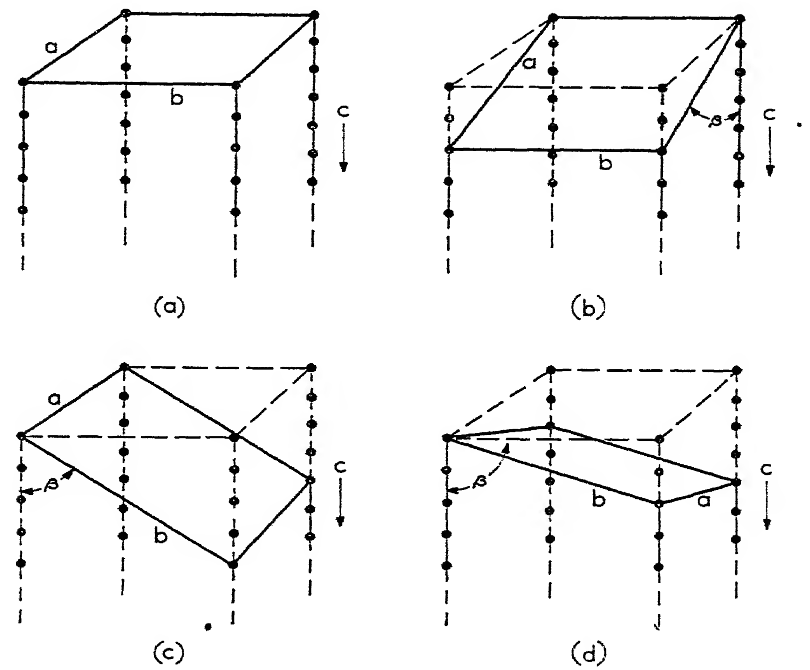


FIG. 2. Origin of polymorphic forms of aliphatic chain compounds (Schoon)

IV. GENERAL CHARACTER OF HIGH MOLECULAR CHAIN STRUCTURES

In the case of the low molecular chain compounds already considered, the crystals are built up of molecules which are identical in structure, size, and shape. These crystals, by virtue of this fact, show well-formed faces and cleavage planes. As the length of the molecules increases, however, the forces arising from the end groups on the molecules become less important in determining the crystal structure than the lateral forces acting between the chains. Under these conditions the simple molecular lattice gives way to what has been called a macromolecular (71) or chain lattice

(*Kettengitter*), in which the ends of the molecules occur at no regular positions in the structure (figure 3), and hence do not give rise to reflections in the x-ray patterns. For convenience, therefore, we may define a high molecular chain system as one which is composed of chain molecules of such length that the effects of the ends of the molecules in determining the structure are negligible in comparison with the rest of the molecule. It follows that x-rays can furnish no direct information regarding the chain length present in these systems.[5] In the ideal case the molecules in such a system are all of the same length. In the case of real systems, however, a distribution of molecule sizes is always present. In this case the definition

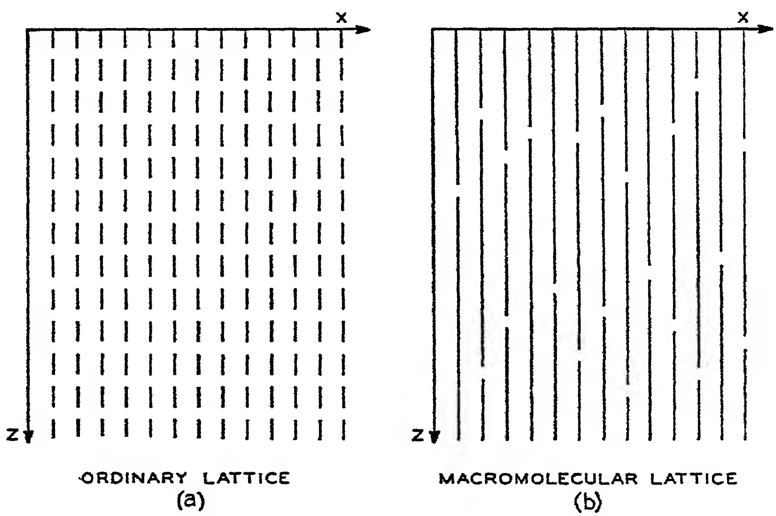


FIG. 3. Plane through (a) an ordinary and (b) a macromolecular type of molecular lattice.

above refers to the average chain length. In general, the chain molecules comprising real high molecular chain systems are produced by the primary valence repetition of a repeating unit or group. It is evident, therefore, from the above definition that the crystal structure in such systems will depend on the nature of this repeating unit and may be described conveniently in terms of it.

Numerous theories regarding the constitution of high molecular substances have been advanced. We will distinguish here only between the theories which view these substances as composed of "micelles" or "crystallites" of more or less definite size and shape (29, 35, 36) and the theories

[5] The proteins are an exception, in that definite units related to the molecule appear to be present (6).

which regard these substances as having a continuous structure containing less well-defined crystalline portions (58, 17, 30, 24).[6] According to the first group of theories the crystalline particles are composed of long chain molecules and are supposed to be held together by micellar forces or by another substance (*Kittsubstanz*). According to the second group the structure is integral in that the chain molecules may traverse both crystalline and amorphous areas. The latter view is taken in this discussion as explaining more satisfactorily the behavior of synthetic high molecular compounds. These compounds may be regarded as consisting of chain molecules of different lengths which, over certain portions of their lengths at least, fit into a lattice-like arrangement (macromolecular lattice). These portions may be termed "crystalline regions" in order to avoid attributing to them any definite form. Over adjacent volumes of the material, however, the chains are to be regarded as in imperfect arrangement and to behave as amorphous or psuedo-crystalline matter. These imperfections in the structure may result from a grouping of molecule ends or from the geometrical inability to fit portions or ends of the long chain molecules into the lattice work. In the case of impure compounds where similar but non-identical chains exist together or where the chain molecule itself possesses an irregular structure, the occurrence of such imperfection areas is obvious (19). The relative proportion of these crystalline and amorphous regions may vary widely. In some cases the amorphous regions may be continuous and in others the crystalline.

On the basis of the above picture we are justified in regarding these compounds as colloidal. It is not correct, however, to regard the system as a crystalline component dispersed in an amorphous one or *vice versa*, since the two regions show no true interfaces but rather a blending of one structure into the other. Regardless of what picture represents the true situation, however, it is only by a consideration of both the amorphous and the crystalline nature of high polymeric substances that a complete explanation of their properties will be obtained. This view has been emphasized by Halle (25) and recently again by Mark (43).

An important characteristic of high molecular chain systems is their property of crystalline orientation. If the crystal regions are not already oriented with respect to one or more directions in the specimen, they are generally susceptible of such orientation by application of the proper stresses. In the case of natural or synthetic fibers the crystalline regions

[6] The distinction made here is for purposes of discussion. Except for differences in detail, the continuous structure point of view is now quite widely held (see Kratky and Mark (37)). Kratky (38), however, has recently presented evidence in favor of discrete crystalline units. Evidence of a continuous net structure for hydrate cellulose has been given by Kratky and Platzek (Kolloid-Z. **88**, 79 (1939)).

generally assume an orientation such that an axis of the crystal lattice lies along the fiber axis but the regions are otherwise random (uniaxial orientation). In all cases of fibers encountered in practice in which uniaxial orientation is present, it has been found that the chain molecules are aligned in the direction of the fiber axis. In films it is sometimes possible to cause the crystal regions to assume a fixed orientation with respect to two directions in the specimen (44, 67). In this case, in effect, we have produced an imperfect single crystal (selective uniaxial orientation) (67). Such foil structures are the most suitable specimens for the determination of the crystal structure of the crystalline regions in the system under investigation.

V. X-RAY RESULTS ON SYNTHETIC POLYMERS

In considering the published data on synthetic polymeric compounds it will be evident that the majority of the investigations is concerned with the structure of the crystalline regions as defined above. Since in many cases this portion represents essentially the whole of the system, this restricted view is generally justified and offers the best approach to the more complex cases where the system contains considerable amorphous matter in addition to the crystalline.

A. Polymethylene oxides

Staudinger (69) was one of the first to appreciate the value of studying synthetic polymers in order to throw light on the nature of high polymeric substance. One of the first series of synthetic compounds to receive attention from both the viscosity and the x-ray standpoints was the polyoxymethylene. Employing carefully fractionated samples of polyoxymethylene diacetates containing from nine to nineteen (CH_2O) repeating units in the molecule, Staudinger, Johner, Signer, Mie, and Hengstenberg (70) showed that each of the lower fractions formed true crystals from which the length of the molecule could be determined. From the uniform increase in this length for the various fractions the length of the CH_2O unit was found to be 1.93 Å. In the case of the higher diacetates, however, the reflections corresponding to the length of the molecule disappeared and a macromolecular lattice (figure 3) was considered to form. Hengstenberg (27) pictured the polyoxymethylene chain as consisting of nine CH_2O units arranged in a threefold screw axis. The fiber period was therefore 17.4 Å. From a high polymer he succeeded in drawing oriented threads which gave a typical x-ray fiber pattern. Ott (55), although agreeing with the Staudinger chain molecule concept, disagreed with the results of Hengstenberg on the formation of a macromolecular lattice. Ott found spacings of 45.1 Å. and 113.4 Å., which he identified with the

molecule lengths in δ- and γ-polyoxymethylene, respectively. Sauter (61) again undertook the study of these compounds but was unable to find the long spacings observed by Ott. From rotation photographs on a minute single crystal of β-polyoxymethylene Sauter determined an identity period of 17.25 Å. along the c-axis (fiber axis) of the crystal. The identity periods in the other two directions were 4.43 Å. and 7.69 Å., in good agreement with the previous work of Hengstenberg. The crystal was believed to belong to the hexagonal space group C_3^2 or C_3^3. In a later paper Sauter (62) considered in detail the chain configuration of the polyoxymethylenes and concluded that the polyoxymethylene chain consists of a nearly planar "tub" form (figure 4b) of chain rather than a zigzag type such as in the

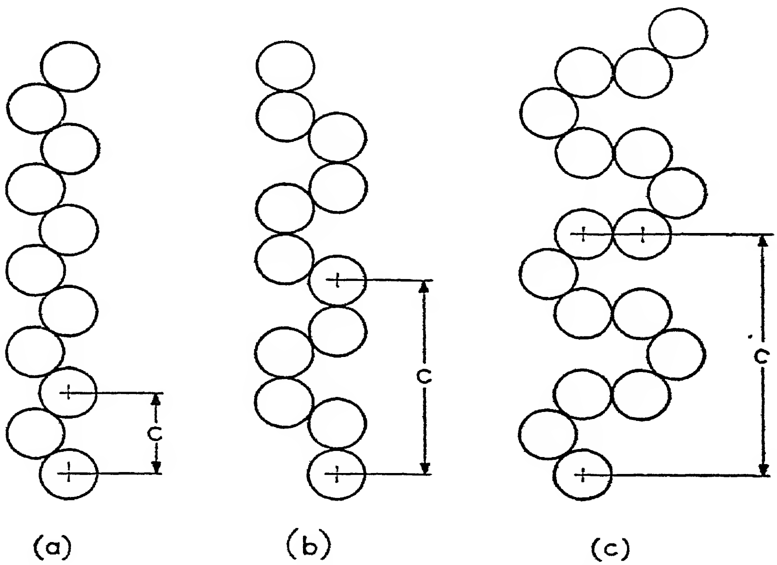


FIG. 4. Planar configurations of single-bonded carbon chains. a, zigzag or *trans* configuration; b, "tub" or *cis* configuration; c, "meander" configuration.

paraffins (figure 4a). The presence of the threefold rhythm, however, necessitates a non-planar arrangement.

In an interesting paper Kohlschütter and Sprenger (34) showed that cyclic trioxymethylene is transformed under the proper conditions into linear polyoxymethylenes. They found that individual fibers gave a highly oriented fiber pattern apparently identical with that previously observed by Hengstenberg.

B. Polyethylene oxides

From a study of polymers obtained by the polymerization of ethylene oxide, Staudinger (69, page 293) concluded that the chain molecules in

solution were much shorter than one would expect from the number of atoms in the molecule. He therefore assumed a "meander form" of chain such as shown in figure 4c. Sauter (63) studied the polyethylene oxides in the solid state and found evidence of a strongly shortened "meander-like" chain as the viscosity measurements indicated for molecules in solution. Sauter found no change in the x-ray patterns (powder) with increase in average molecular weight. From an elongated sample of high polymeric polyethylene oxide he obtained a fiber pattern from which a fiber period of 19.5 Å. was calculated. The unit cell is bounded by four chain molecules aligned in the fiber direction and perpendicular to a simple monoclinic base having a = 9.5 Å., c = 12.0 Å., and β = 101°. Thirty-six ethylene oxide units are present in the cell. From symmetry and bond considerations Sauter deduced that nine ethylene oxide units were present in each chain over the distance corresponding to the fiber period. This requires a highly folded configuration and Sauter was led to regard it as a meander-like chain, although he did not picture it as being necessarily planar. H. Staudinger, M. Staudinger, and E. Sauter (72) recently have summarized their findings on the crystalline nature of the polyethylene oxides and polymethylene oxides. They obtained microscopic evidence showing that the crystal size in these polymers varies inversely as the degree of polymerization. For degrees of polymerization above 500 they found that the polyethylene oxides could be "cold-formed" into fibrillar regions of definite cross section. Fibrils of 0.15 μ (1500 Å.) diameter were reported in the case of the polyoxymethylenes from examinations under the ultraviolet microscope. By way of comparison they suggest that fibrillar aggregates as low as 100 Å. in diameter are probably present in cellulose because of fibrillar shattering. The examination of single crystals of β-polyoxymethylene showed no basal cleavage but a rough surface fracture. Bending of these crystals caused fibrillar shattering, indicating them to be intrinsically brittle.

The main findings of Staudinger and his associates on the polymethylene and polyethylene oxides may be summarized as follows: (*1*) As in the case of the low molecular chain compounds, the molecules in the synthetic polymers crystallize with the chains parallel. Oriented fibers of the polymers, like oriented natural fibers, contain the long axis of the chain molecules parallel to the fiber axis[7]. The configuration of the long chain molecules, however, is variable and dependent on the nature of the chemical repeating unit. (*2*) It is unnecessary for the molecules of a polymeric series to be of equal length (molecular weight) in order to form crystals. In such macromolecular crystals the end groups represent imperfections

[7] Generally the molecules are included in a narrow cone, the fiber axis forming the axis of the cone.

in the lattice (figure 3), and the chemical repeating unit in the polymer may be considered as the unit of packing. A macromolecular lattice is also supposed to form in the case of equally long molecules of sufficiently great length. There may be, however, in view of Ott's findings, some question as to whether this latter point has been sufficiently proven, since it is extremely difficult to prepare high polymers of strictly uniform length.

Barnes and Ross (3) also studied the polyethylene oxides and compared their x-ray patterns with the higher polyethylene glycols. They were able to detect no difference in the Debye–Scherrer patterns of the polymers prepared by the different methods.

C. Linear polyesters

As a result of previous work by Carothers (9, 10), Carothers and Hill (11) in 1932 prepared linear polymers by the intermolecular polycondensation of dibasic acids and glycols which exhibited the remarkable property of orienting or "cold drawing" when stress was applied to unoriented rods of the material. A. W. Kenney (11) showed that cold-drawn fibers of polyethylene sebacate gave well-oriented fiber patterns from which identity periods could be calculated corresponding roughly to the length of the repeating unit in the chain molecules.

Fuller and Erickson (20) later undertook an x-ray study of various linear polyesters. They found that, analogous to the polymethylene and polyethylene oxides, the polyesters showed identical Debye–Scherrer patterns with increasing average molecular weight. Oriented (uniaxial) specimens of high polymeric ethylene succinate, ethylene adipate, ethylene azelate, ethylene sebacate, trimethylene sebacate, and diethylene sebacate were examined. The identity periods along the fiber axis were obtained from the fiber patterns of the various compounds. It was concluded from this work that in all cases the chain molecules are aligned along the fiber direction. Comparison of calculated lengths for various chain configurations shows that the chain molecules in the adipic, azelaic, and sebacic esters of ethylene glycol agree well with a planar zigzag paraffin type of chain. The succinate ester requires an entirely different chain configuration from the other compounds of the series. A helical type of chain somewhat analogous to that observed by Sauter (62) for polyoxymethylene was found to agree best with the results. The chains in the trimethylene and diethylene polyesters of sebacic acid also deviated from the zigzag form and probably represent helical or folded structures. Storks (73) applied the electron diffraction method successfully to thin films of polyethylene succinate, adipate, and sebacate. His results on oriented specimens agree well with those of Fuller and Erickson (20). Storks concluded that the crystals present in these films are of very small cross section, since they

are randomly rotated in the cross section of films less than 400 to 1000 Å. thick. The fiber patterns also clearly showed the presence of meridian reflections which (because of the diffraction angle) were not observed in the case of the x-ray patterns. These reflections in the case of the succinate indicate intermediate planes at 2.1 Å. along the fiber axis and in the case of the other two esters confirm the correctness of the zigzag type of chain configuration. Finally, Storks observed selective uniplanar orientation (67) in unstretched films of polyethylene sebacate. In this case the crystals, although random in the film plane, show an approximately fixed orientation about the chain axis. The work of Storks illustrates clearly the utility of electron methods in the study of thin films of organic polymers.

Subsequently, Fuller and Frosch (21) summarized the work on the ethylene series of polyesters and in a later paper (22) treated the decamethylene series of polyesters. The polymeric self-ester of hydroxydecanoic acid was also considered in the former paper. Without going into detail, the results of these studies may be summarized as follows:

The ethylene series of polyesters above the glutarate and the decamethylene series above the carbonate are found to possess chains which are essentially planar zigzag in configuration. Figure 5 shows the regular increase in the length of the repeating unit with the number of chain atoms for each series. The slope is the same in both cases and corresponds to 1.26 Å. per CH_2 group. This is precisely the slope calculated on the basis of a tetrahedral arrangement of carbon atoms at a distance of 1.54 Å. from each other (51). The slight displacement of the curves for the two series from one another may be real (20). It is, however, close to the experimental error. The deviation from the calculated curve may be due to an error in the assumed oxygen valence angle or in the C—O bond distance (74). The chain molecules pack in the crystal in the same manner as has been found to hold for the paraffins (51, 60)[8]. There is some evidence that a non-base-centered arrangement, such as Kohlhaas (32) has proposed for cetyl palmitate, is the correct one. We may picture the structure of these compounds as shown in figure 6a, which shows the type of chain arrangement assumed in the case of ethylene azelate and in general by the polyesters which contain an odd number of chain atoms. In this case the carbonyl groups in adjacent chains align themselves in horizontal planes. The chains in these esters (as can be seen from their structural formula) possess a twofold screw axis. Figure 6b shows the situation which holds for most of the even esters of both series. In this case, although the chains pack in cross section as in figure 6a, they are no longer arranged with the

[8] Decamethylene oxalate is an exception. In this case, the cross section of the unit cell is broader one way and narrower the other than in the case of the paraffins.

carbonyl groups opposite but are apparently displaced various amounts parallel to the chains for different even esters. For comparison the fiber patterns observed for polyethylene azelate (odd) and polyethylene sebacate (even) are given in figures 7a and 7b, respectively. The ethylene esters from the adipate on all appear to conform to the same pattern:—the chains are displaced parallel to their length a distance of two chain atoms

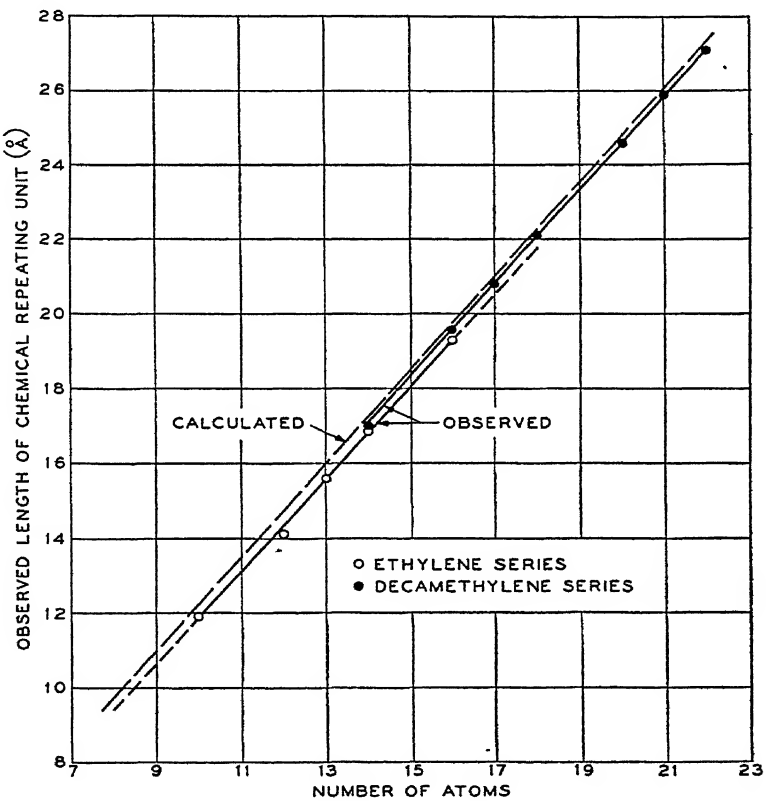


FIG. 5. Length of chemical repeating unit as a function of number of chain atoms for the ethylene and decamethylene series of polyesters.

in successive (100) planes. This arrangement is analogous to that shown in figure 2b. In the case of the decamethylene esters, with the exception of the oxalate which packs in cross section somewhat differently from the others, the situation is similar. However, instead of a constant displacement of successive planes a distance of two chain atoms, as in the case of the even members of the ethylene series, we find evidence of glides of other magnitudes. Thus, decamethylene sebacate appears to be analogous to

the ethylene case, whereas decamethylene succinate agrees best with a displacement of four chain atoms in (100) or five chain atoms in (010). In decamethylene suberate half of the chain molecules appear to have been

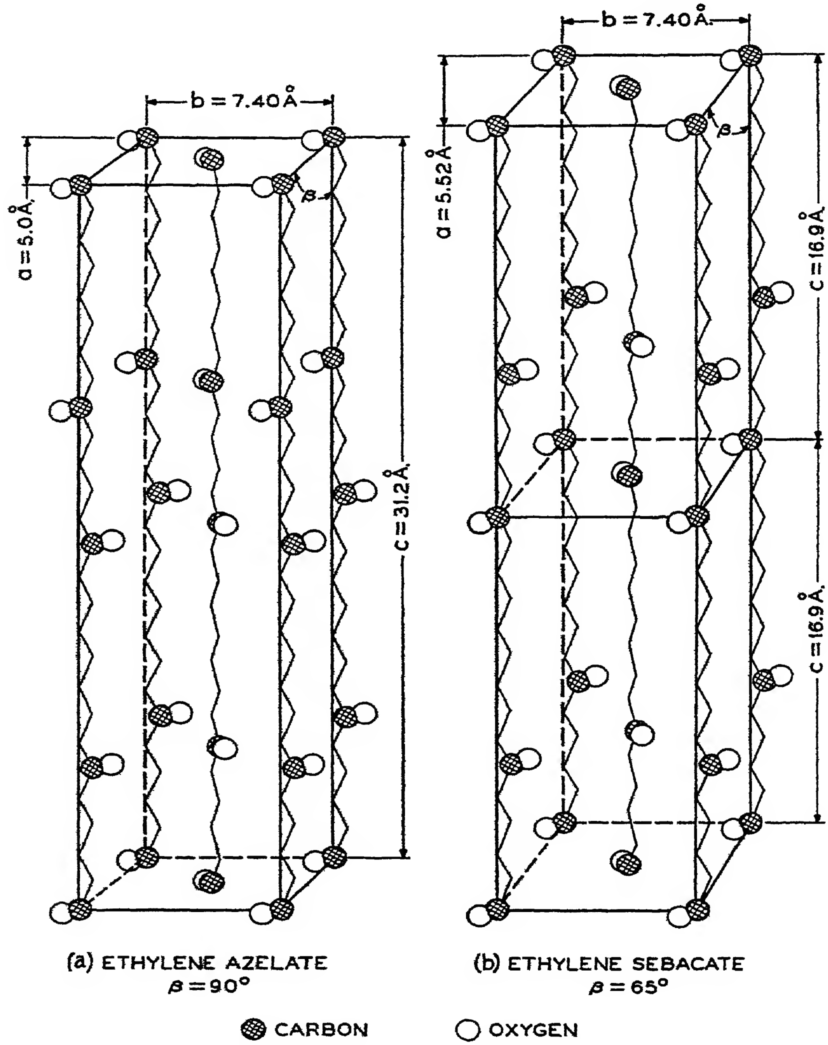


FIG. 6. a, orthorhombic unit cell of polyethylene azelate; b, monoclinic unit cell of polyethylene sebacate, $\beta = 65°$. Only the carbonyl groups are represented.

displaced one-half a repeating distance relative to the other half. In the case of the odd decamethylene esters, as in the odd ethylene esters, second-order meridian reflections are observed, indicating a chain alignment in

the crystal as shown in figure 6a. The situation, however, is complicated by the presence of "even" type reflections as well. Likewise, the even adipate ester is complicated by the presence of "odd" type reflections.

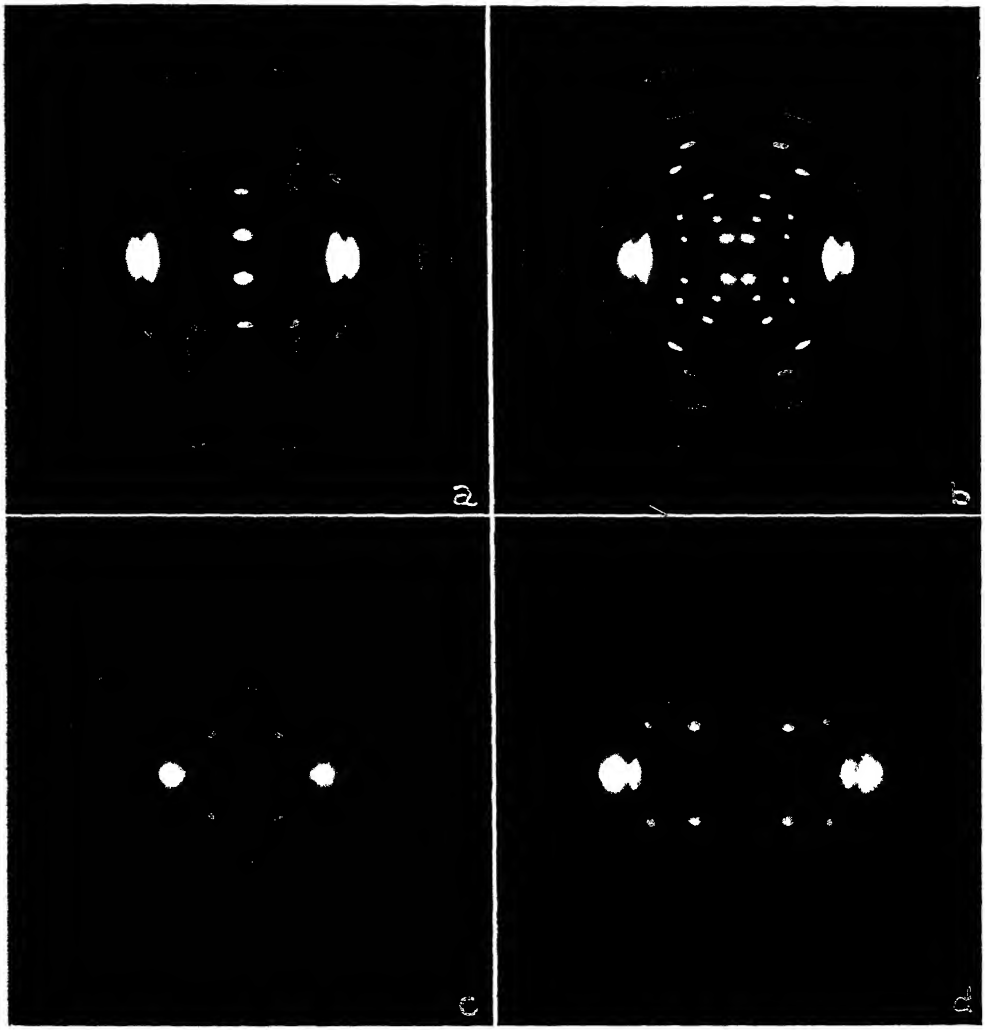


FIG. 7. X-ray fiber patterns; fiber axis vertical. a, polyethylene azelate, 3.5 cm.; b, polyethylene sebacate, 3.5 cm.; c, natural silk, 3.0 cm.; d, polyhexamethylene adipamide, 4.0 cm.; e, polyethylene tetrasulfide, 2.5 cm.: f, polyethylene disulfide, 3.5 cm.; g, gutta-percha, β-form, 4 0 cm ; h, polychloroprene, 4.0 cm.; i, polyvinyl alcohol, 4.0 cm.; j. polyvinyl chloride, 4.0 cm ; k, polyvinylidene chloride, 4.0 cm.; m, polyisobutylene, 4.0 cm.

The best explanation of the simultaneous occurrence of these reflections at the present time is that in these members of the decamethylene series, crystals of both the "odd" (figure 6a) and "even" (figure 6b) types are present.

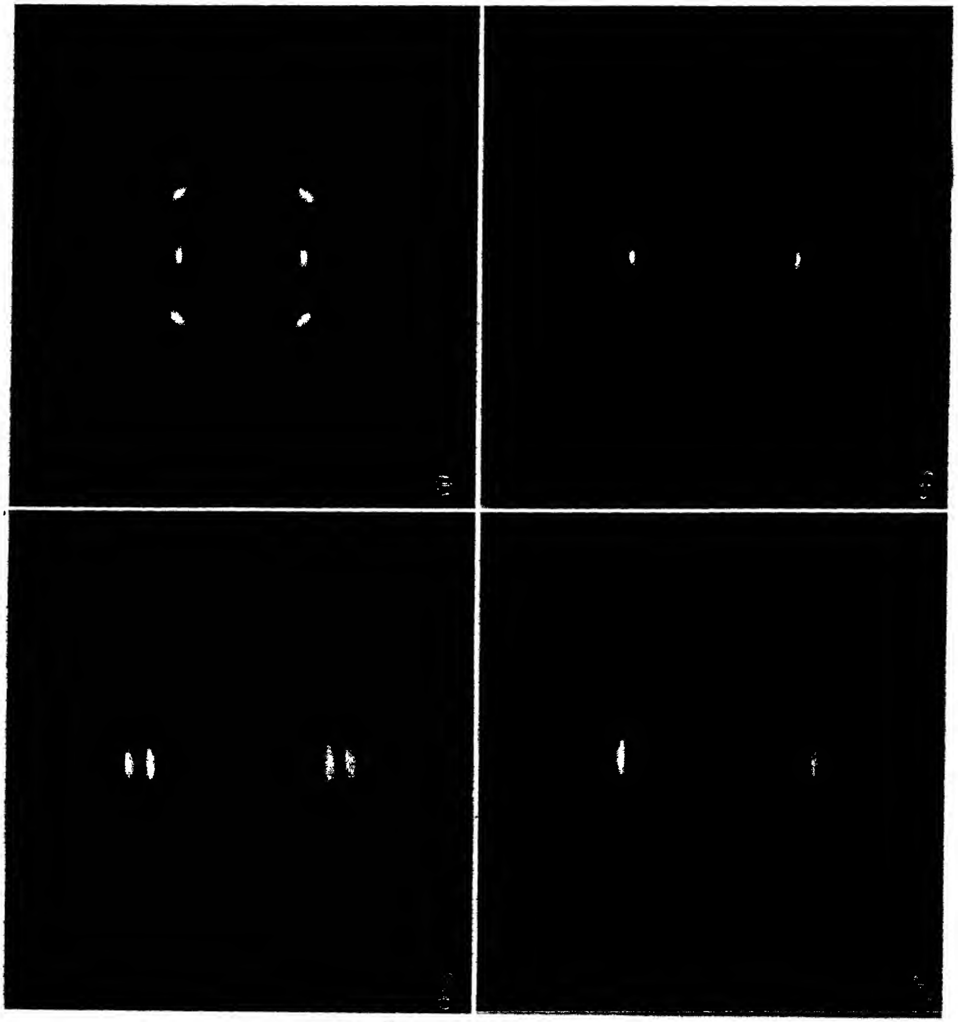


FIG. 7. e, f, g, h

The analogy of the above results to those reported for the low molecular chain compounds is immediately evident. The orthorhombic and monoclinic forms reported for the paraffins and fatty acids (figures 2a, 2b, and

2c) correspond to the forms of the esters shown in figures 6a and 6b, respectively. Thus the structure of the even ethylene esters corresponds closely to the monoclinic form for stearic acid (a = 5.546 Å., b = 7.38 Å,

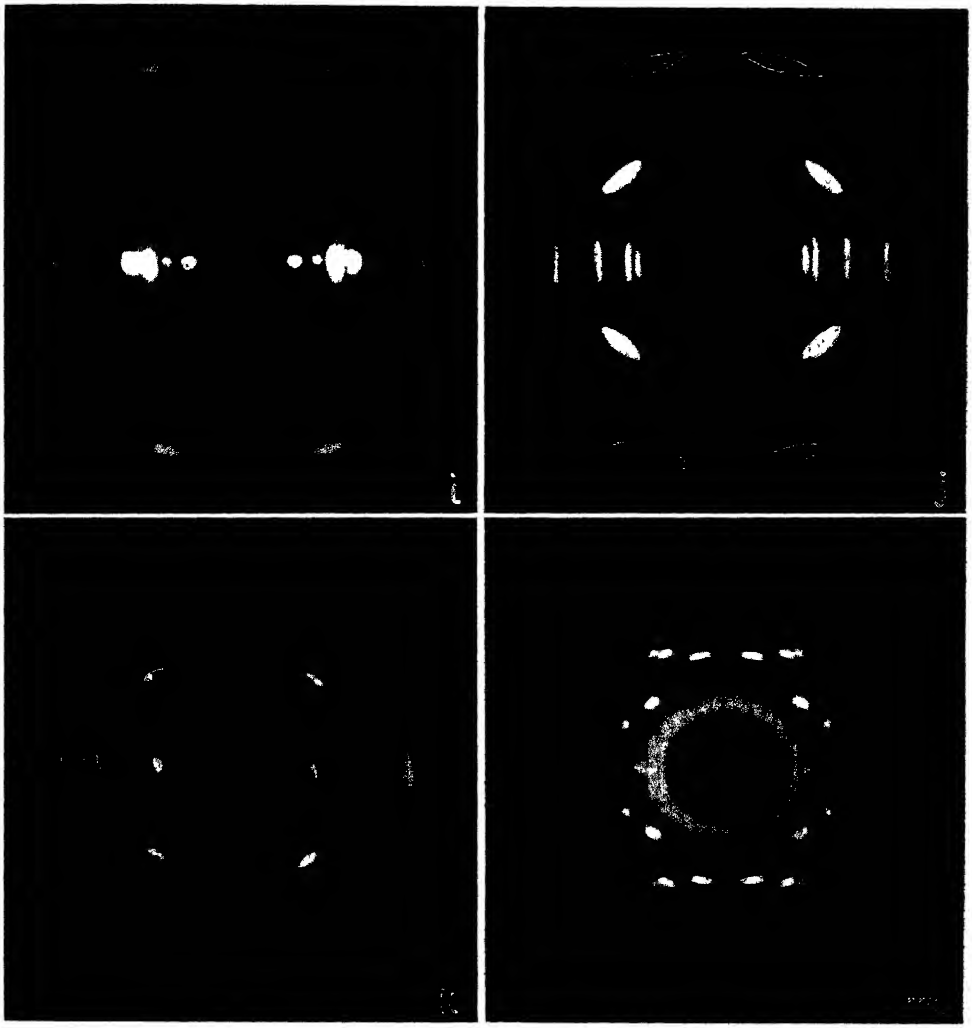


FIG. 7. i, j, k, m

and β = 63° 38′ reported by Müller (51)) except for the c-edge of the cell[9]. The applicability of the theory of Schoon (65) to these polymeric com-

[9] A triclinic cell for these esters is not excluded by the present data, but in view of the results on the aliphatic compounds this cell appears unlikely.

pounds is therefore evident. Just as in the case of the low molecular compounds, we may assume a molecule gliding, resulting in a variety of stable crystal forms. The "tilt" of the chains in the low molecular compounds corresponds to a uniform gliding of the molecules along the fiber axis in the case of the high molecular compounds. So far, however, it has not been possible to bring about changes in these forms by heating at definite temperatures, as has been done in the low molecular crystals.

D. Linear polyamides

No work relating to the crystalline nature of the polyamides (12) has yet appeared. Unpublished data of the author, however, show a close analogy to exist between these chain compounds and the linear polyesters. Thus, figure 7c shows a fiber pattern of polyhexamethylene adipamide[10]. The observed fiber period of 17.0 Å. agrees well with that calculated (17.3 Å.) on the basis of a planar zigzag chain[11]. A fiber pattern of natural silk is given in figure 7c for comparison. The pattern (figure 7d) shows that essentially planar zigzag chains are present in somewhat different arrangement in cross section from that in the paraffins. The number of possible polyamide structures is very great, and this class of linear polymers will in the future undoubtedly furnish most valuable information on the structure of both synthetic and natural high polymers.

E. Linear polysulfides

Another interesting class of synthetic compounds on which little has been published is the polymeric organic polysulfides (41). Katz became interested in these compounds in 1934. Katz and Fuller (31) showed certain of them to be crystalline in the unstretched condition and to be capable of a high degree of orientation on elongation. Figure 7e shows the fiber pattern produced by the compound formed by the reaction of ethylene dichloride and sodium tetrasulfide.

The pattern appears to be a composite one, consisting of a strong fiber diagram superimposed on a weaker one. Whether the weaker pattern is due to an impurity (57) or to another crystal modification of the tetrasulfide is not known. Neglecting the reflections which do not rationalize with the main layer-line system, we obtain a fiber period of 4.32 ± 0.03 A. (31). This value agrees with a type of chain (not necessarily planar) shown in figure 8d, in which two sulfur atoms are joined to the ethylene residue and two are joined to these sulfurs to form a ring[12]. This picture

[10] This material was supplied by the du Pont Company.

[11] Brockway's value of 1.46 Å. for C—N (Rev. Modern Phys. **8**, 231 (1936)) is taken. The angles are assumed tetrahedral.

[12] Figure 8 is schematic. It is intended to represent structures which are possible and which agree with the present x-ray data.

appears more likely than the one previously given (31), in which two sulfurs were attached latterly to two sulfurs in the chain. As Patrick (57) has shown, there is good chemical evidence for the fact that two of the sulfur atoms are bonded differently than the other two.

Unlike polyethylene tetrasulfide, polyethylene disulfide[13] prepared from ethylene dichloride and sodium sulfide gives a fiber pattern (figure 7f)

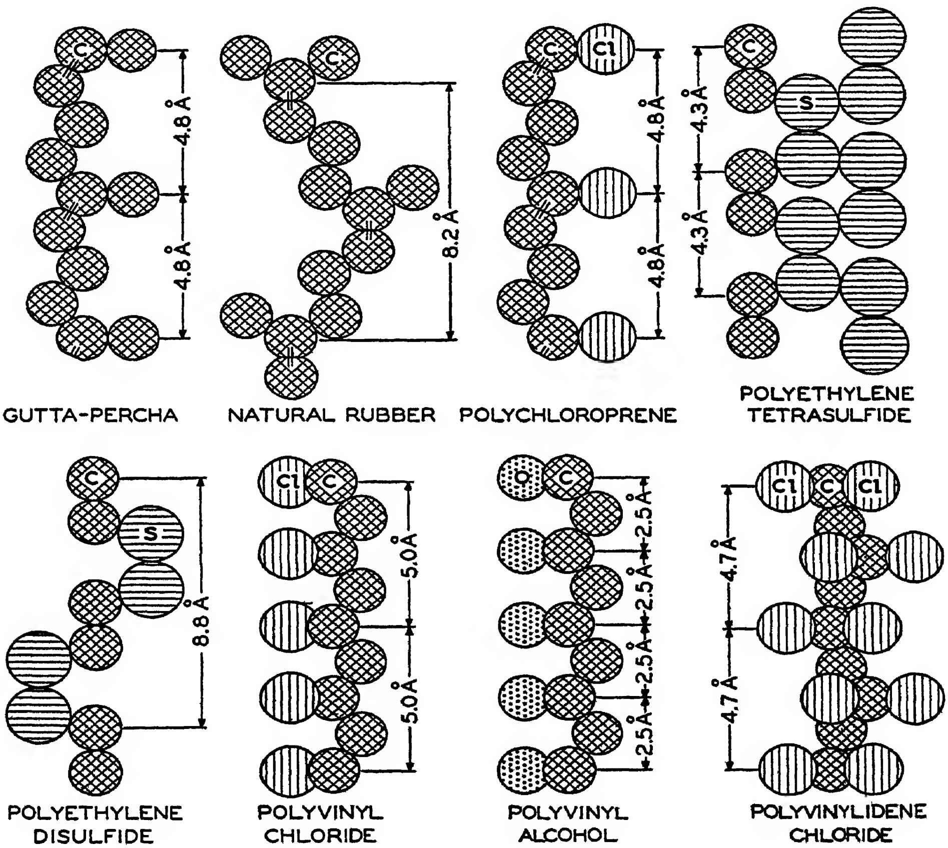


FIG. 8. Chain configurations of various linear polymers

which is indicative of a pure compound. The observed fiber period is 8.8 ± 0.05 Å. The appearance of a strong second-order meridian reflection, together with a fiber period which is almost exactly two times that observed for the tetrasulfide, strongly suggests that the disulfide possesses a chain configuration like that shown in figure 8e, in which every alternate

[13] These results are from unpublished work which the author did in collaboration with J. R. Katz. Thanks are due Dr. J. C. Patrick for supplying this sample.

repeating $C_2H_4S_2$ group is rotated through 180°. The length of this unit of between 4.3 and 4.4 Å. is in good agreement with the value 4.34 Å., calculated from the planar form shown in figure 8d assuming C—C = 1.52 Å., C—S = 1.75 Å., and angles of 109.5° and 103° (56) at carbon and sulfur, respectively.

F. Vinyl derivatives

The synthetic polymers produced by the polymerization of compounds possessing the vinyl group, although of great industrial importance, have not been the subject of much careful x-ray investigation. These compounds are now generally conceded to be formed as a result of an activated chain reaction, which results in a distribution of polymers of various chain lengths. Evidence has been presented by Marvel (45, 46, 47) for a "head-to-tail" arrangement of repeating units in most of these compounds. Because of the nature of the polymerization reaction, however, the chain molecules do not appear to possess as high a degree of linearity and regularity as those produced by polycondensation. Nevertheless, in many cases a considerable degree of crystallinity is present and useful information can be obtained by x-ray study.

(*1*) *Polyethylene.* By the polymerization of ethylene it has been possible (16) to build extremely long hydrocarbon chains. Because of the twofold symmetry of the repeating unit, regular molecules arise in this case which crystallize in a macromolecular type of lattice. In a recent work Bunn (8) has determined that the repeating units, as might be anticipated, are arranged in an orthorhombic cell (*Pnam*) having a = 7.4 Å., b = 4.93 Å., and c = 2.534 Å. The chains are paraffinic in configuration (figure 4a) and in a rolled sheet are aligned in the rolling direction with the (110) plane in the sheet plane. From a comparison of the observed and calculated intensities Bunn concludes that the spherical carbon atom in the present conception of the hydrocarbon chain must be replaced by a carbon atom with an electron cloud flattened in the direction of the chain. Presumably this finding applies to all primary valence chains.

(*2*) *Polyhaloprenes.* Carothers and Kenney (13, 14) showed that chloroprene and bromoprene are amorphous in the unstretched condition but crystallize on elongation, giving a fiber diagram as in the case of natural rubber. The observed fiber period is 4.8 Å. in the case of chloroprene.[14] The authors point out that the chain molecules approximate the gutta-percha type of configuration rather than the more folded rubber type. The close analogy to gutta-percha is shown by their respective fiber patterns in figures 7g and 7h. Garbsch and v. Susich (23) obtained a fiber

[14] Unpublished work of the author shows polybromoprene to have a fiber period practically identical with that of polychloroprene.

period of 4.8 ± 0.1 Å. at 600 per cent elongation on chloroprene rubber. Krylov (39) studied chloroprene with electron rays and reported an orthorhombic cell with a = 9.0 Å., b = 8.23 Å., and c = 4.79 Å. (fiber axis). Both chloroprene and bromoprene give broad fuzzy reflections in the fiber pattern. This is indicative of an imperfect arrangement of repeating units in the lattice (perhaps because of irregular chain structure) or of extremely thin crystals (64). A comparison of the chain structure for chloroprene with that for gutta-percha is shown in figures 8a and 8c. Since the calculated fiber period for these configurations is approximately 0.2 Å larger than the observed, it is probable that the chain deviates from planar. The Meyer and Mark (48) structure for natural rubber is given in figure 8b.

(*3*) *Polyvinyl alcohol.* Halle and Hofmann (26) have obtained a sharp fiber diagram from polyvinyl alcohol by stretching under the proper conditions. They find a fiber period of 2.57 Å. This is slightly more than the value expected from a planar zigzag chain, and the authors suggest that this may be due to an increase in the tetrahedral valence angle.

The author[15] has repeated this work recently and has found from a sharp fiber pattern (figure 7i) an identity period of 2.52 ± 0.02 Å., in very good agreement with the accepted value for the alternate carbon-to-carbon distance along a planar zigzag chain. The configuration of the polyvinyl alcohol chain is shown in figure 8g and conforms to the finding of the "head-to-tail" structure in this compound by Marvel and Denoon (46).

(*4*) *Polyisobutylene.* Brill and Halle (7) examined polyisobutylene by means of x-rays. They found it to give on stretching a fiber pattern similar to that of natural rubber. The fiber period observed was 18.5 Å. They concluded that helical rather than zigzag chains are present. A fiber pattern obtained on this material[16] by the author is shown in figure 7m. An extremely high degree of orientation and crystallinity is indicated. The meridian reflection observed on the eighth layer-line arises from intermediate planes perpendicular to the fiber axis and 2.3 Å. apart. A "head-to-tail" structure is therefore indicated, the shortening arising from a helical arrangement of the pairs of methyl groups such as Brill and Halle have suggested.

(*5*) *Miscellaneous vinyl derivatives.* Natta and Rigamonti (54) examined a number of vinyl polymers by means of electron rays. They found polyvinyl acetate to be amorphous, polyvinyl chloride and polyvinyl bromide to be partly crystalline, and polyvinylidene chloride to be quite highly crystalline. Misch and Picken (49), however, found a thick-

[15] Unpublished work.

[16] The author is indebted to Mr. F. J. Malm for this sample.

ening of the amorphous ring of polyvinyl acetate on stretching at 60°C., in agreement with Katz (31), and concluded that irregular chains are probably present.

The author has found a fiber period[17] of 5.0 ± 0.05 Å. for polyvinyl chloride (figure 7j). This value indicates the presence of a planar zigzag chain and in view of the proof of the 1:3 structure of vinyl chloride (45) necessitates that every other chlorine atom be differently disposed along the chain. A suggested arrangement is that shown in figure 8f, in which the chlorines take up alternate positions on opposite sides of the chain. Such a configuration is in agreement with the repulsion of chlorine atoms at this distance, as found by Beach and Palmer (4) in the case of ethylene dichloride.

The repulsion between chlorine atoms apparently also plays a part in determining the structure of polyvinylidene chloride, $(—CH_2—CCl_2—)_n$. Oriented samples of this substance give sharp fiber diagrams (figure 7k). The observed fiber period[18] is 4.7 ± 0.05 Å. A strong second-order meridian reflection is also present, which indicates a 1:3 disposition of the pairs of chlorine atoms. The fiber period shows a definite shortening compared to the planar zigzag arrangement of carbon atoms which requires a period of 5.0 Å. Such a shortening can be obtained by a partial rotation of alternate pairs of chlorine atoms, as illustrated in figure 8h. It is evident that this arrangement, with the individual chlorine atom the maximum distance from its neighbors, is in agreement with the repulsion of these atoms mentioned above.[19]

There can be little doubt that in the case of many of the vinyl polymers we have to do with systems containing much amorphous material. Katz (31) has pointed out that in many cases the "amorphous" rings in these substances correspond closely to rings given by the liquid monomers themselves, and that they often give in addition another ring which he calls the "polymerization ring." Thus, in the case of polystyrene a ring (d = 4.8 Å.) corresponding to liquid styrene and an inner ring (d = 10 Å.) corresponding to an interchain spacing are observed. Likewise in polyvinyl acetate an inner or polymerization ring (d = 7.0 Å.) is observed, which Katz identifies with the interchain distance. In both polystyrene and polyvinyl acetate (as mentioned above) it has been possible to cause

[17] The pattern was not sufficiently well formed to exclude the fiber period being a multiple of 5.0 Å. The pattern shown in figure 7j has been reconstructed from the original and shows the positions of the main reflections.

[18] Unpublished work of the author.

[19] An alternative configuration for polyvinylidene chloride which is not excluded by the present evidence is a helical arrangement such as has been suggested for polyisobutylene. The fiber period would then be a multiple of 4.7 Å.

a splitting of these inner rings corresponding to a parallel ordering of the chains. There is therefore good evidence that the long chains in many of these polymers exist in a sort of mesomorphic arrangement, in which portions of the chains are arranged parallel and at a given distance apart but are otherwise unordered.

VI. CONCLUSION

The ultimate aim of research on high polymeric substances is to explain the properties of these substances in terms of their inner colloidal and molecular nature. X-ray and electron diffraction investigations, as we have seen, have furnished considerable data on this inner structure. Nevertheless, it is of the utmost importance that the investigator in this field does not obtain a one-sided view from the past work. Nearly always a given study relates to only one phase of the general problem. For example, it is wrong to assume that, since the "unit cell" for cellulose has been well established, we therefore know its entire structure. In reality we have determined simply that parts of the cellulose have this structure. How these parts fit into the other parts that make up the entire system is still not clear. Here again the synthetic high polymers can be of great help. By employing simple polymeric molecules of known chemical constitution and structure and by determining how these molecules behave in the aggregate, we may expect to make progress in understanding these complex systems. In this endeavor a closer control of both the specimens under examination and the technique itself will serve to make the diffraction method an even more valuable tool for the study of high molecular compounds.

The author is indebted to Mr. C. J. Frosch and Dr. W. O. Baker for many helpful suggestions. Thanks are also due to Mr. N. R. Pape for aid in obtaining certain of the x-ray results.

REFERENCES

(1) Astbury, W. T.: *Fundamentals of Fibre Structure.* Oxford University Press, London (1933).
(2) Astbury, W. T.: Proc. Roy. Soc. (London) **A150**, 533 (1935).
(3) Barnes, W. H., and Ross, S.: J. Am. Chem. Soc. **58**, 1129 (1936).
(4) Beach, J. Y., and Palmer, K. J.: J. Chem. Phys. **6**, 639 (1938).
(5) Bernal, J. D.: Z. Krist. **83**, 153 (1932).
(6) Bernal, J. D.: Nature **143**, 663 (1939).
(7) Brill, R., and Halle, F.: Naturwissenschaften **26**, 12 (1938).
(8) Bunn, C. W.: Trans. Faraday Soc. **35**, 483 (1939).
(9) Carothers, W. H.: J. Am. Chem. Soc. **51**, 2548 (1929), *et seq.*
(10) Carothers, W. H.: Chem. Rev. **8**, 353 (1931).
(11) Carothers, W. H., and Hill, J. W.: J. Am. Chem. Soc. **54**, 1579 (1932).

(12) Carothers, W. H., and Hill, J. W.: J. Am. Chem. Soc. **54**, 1568 (1932). Carothers, W. H.: U. S. patent 2,130,948.
(13) Carothers, W. H., Kirby, J. E., and Collins, A. M.: J. Am. Chem. Soc. **55**, 789 (1933).
(14) Carothers, W. H., Williams, I., Collins, A. M., and Kirby, J. E.: J. Am. Chem. Soc. **53**, 4203 (1931).
(15) Clark, G. L.: *Applied X-Rays*, 2nd edition. McGraw-Hill Book Company, Inc., New York (1933).
(16) Fawcett, E. W., *et al.*: British patent 471,590.
(17) Frey-Wyssling, A.: Protoplasma **25**, 261 (1936); **27**, 372 (1937).
(18) Frumkin, L. S., and Goncharov, I. D.: Zarodskoya Lab. **3**, 1112 (1934).
(19) Fuller, C. S.: Ind. Eng. Chem. **30**, 472 (1938).
(20) Fuller, C. S., and Erickson, C. L.: J. Am. Chem. Soc. **59**, 344 (1937).
(21) Fuller, C. S., and Frosch, C. J.: J. Phys. Chem. **43**, 323 (1939).
(22) Fuller, C. S., and Frosch, C. J.: J. Am. Chem. Soc. **61**, 2575 (1939).
(23) Garbsch, P., and v. Susich, G.: Kautschuk **8**, 122 (1932); Rubber Chem. Tech. **6**, 113 (1933).
(24) Gerngross, O., and Herrmann, K.: Z. physik. Chem. **B10**, 371 (1930); Kolloid-Z. **60**, 276 (1936).
(25) Halle, F.: Kolloid-Z. **69**, 324 (1934).
(26) Halle, F., and Hofmann, W.: Naturwissenschaften **23**, 770 (1935).
(27) Hengstenberg, J.: Ann. Physik **84**, 245 (1927).
(28) Hengstenberg, J.: Z. Krist. **67**, 583 (1928).
(29) Herzog, R. O.: Kolloid-Z. **39**, 98 (1926).
(30) Katz, J. R.: *Die Röntgenspektrographie als Untersuchungsmethode.* Urban & Schwarzenberg, Berlin (1934).
(31) Katz, J. R.: Trans. Faraday Soc. **32**, 77 (1936).
(32) Kohlhaas, R.: Z. Krist. **98**, 418 (1938).
(33) Kohlhaas, R., and Soremba, K. H.: Z. Krist. **100**, 47 (1938).
(34) Kohlschütter, H. W., and Sprenger, L.: Z. physik. Chem. **B16**, 284 (1932).
(35) Kratky, O.: Kolloid-Z. **70**, 14 (1935).
(36) Kratky, O.: Naturwissenschaften **26**, 94 (1938).
(37) Kratky, O., and Mark, H.: Z. physik. Chem. **B36**, 129 (1937).
(38) Kratky, O., and Schossberger, F.: Z. physik. Chem. **B39**, 145 (1938).
(39) Krylov, K. I.: Caoutchouc and Rubber (U. S. S. R.), No. 1, 24 (1938); No. 2, 26 (1938); J. Exptl. Theoret. Phys. (U. S. S. R.) **5**, 524 (1935); Rubber Chem. Tech. **9**, 409 (1936).
(40) Langmuir, I.: J. Am. Chem. Soc. **39**, 1848 (1917).
(41) Martin, S. M., Jr., and Patrick, J. C.: Ind. Eng. Chem. **28**, 1145 (1936).
(42) Mark, H.: *Physik und Chemie der Cellulose.* J. Springer, Berlin (1932).
(43) Mark, H.: Paper presented at the Ninety-seventh Meeting of the American Chemical Society, held in Baltimore, Maryland, March, 1939.
(44) Mark, H., and v. Susich, G.: Kolloid-Z. **46**, 11 (1928).
(45) Marvel, C. S.: J. Am. Chem. Soc. **61**, 3241 (1939).
(46) Marvel, C. S., and Denoon, C. E., Jr.: J. Am. Chem. Soc. **60**, 1045 (1938).
(47) Marvel, C. S., and Levesque, C. L.: J. Am. Chem. Soc. **60**, 280 (1938); **61**, 1682 (1939).
(48) Meyer, and Mark, H.: *Der Aufbau der hochpolymeren organischen Naturstoffe.* Akademische Verlagsgesellschaft, Leipzig (1930).
(49) Misch, L., and Picken, L.: Z. physik. Chem. **B36**, 398 (1937).
(50) Müller, A.: J. Chem. Soc. **123**, 2043 (1923).

(51) Müller, A.: Proc. Roy. Soc. (London) **A120**, 437 (1928).
(52) Müller, A., and Shearer, G.: J. Chem. Soc. **123**, 3156 (1923).
(53) Natta, G.: IX Congr. intern quím. pura aplicada **4**, 208 (1934).
(54) Natta, G., and Rigamonti, R.: Atti. accad. Lincei **24**, 381 (1936).
(55) Ott, E : Science **71**, 465 (1930).
(56) Palmer, K. J.: J. Am. Chem. Soc. **60**, 2360 (1938).
(57) Patrick, J. C.: Trans. Faraday Soc. **32**, 347 (1936).
(58) Pierce, F. T.: Trans. Faraday Soc. **26**, 809 (1930).
(59) Piper, S. H , and Grindley, E.: Proc. Phys. Soc. (London) **35**, 269 (1923); **36**, 31 (1923).
(60) Rigamonti, R.: Gazz. chim. ital. **66**, 174 (1936).
(61) Sauter, E.: Z. physik. Chem. **B18**, 417 (1932).
(62) Sauter, E.: Z. physik. Chem. **B21**, 186 (1933).
(63) Sauter, E.: Z. physik. Chem. **B21**, 161 (1933).
(64) Schiebold, E.: Kolloid-Z. **69**, 266 (1934).
(65) Schoon, T.: Z. physik. Chem. **B39**, 385 (1938).
(66) Shearer, G.: J. Chem. Soc. **123**, 3152 (1923).
(67) Sisson, W. A.: J. Phys. Chem. **40**, 343 (1936).
(68) Sisson, W. A.: Ind. Eng. Chem. **30**, 530 (1938).
(69) Staudinger, H.: *Der Aufbau der hochmolekularen organischen Verbindungen.* Hirschwaldsche Buchhandlung, Berlin (1932).
(70) Staudinger, H., Johner, H., Signer, R., Mie, G., and Hengstenberg, J.: Z. physik. Chem. **A126**, 425 (1927).
(71) Staudinger, H., and Signer, R.: Z. Krist. **70**, 193, 202 (1929).
(72) Staudinger, H., Staudinger, M., and Sauter, E.: Z. physik. Chem. **B37**, 403 (1937).
(73) Storks, K. H.: J. Am. Chem. Soc. **60**, 1753 (1938).
(74) Sutton, L. E., and Brockway, L. O.: J. Am. Chem. Soc. **57**, 473 (1935).
(75) Thiessen, P. A., and Schoon, T.: Z. physik. Chem. **B36**, 195, 216 (1937).
(76) Thiessen, P. A., and Spychalski: Z. physik. Chem. **A156**, 309, 435 (1931).

X-RAY INVESTIGATIONS OF CARBOHYDRATES[1]

H. MARK

Canadian International Paper Co., Hawkesbury, Ontario, Canada

Received October 12, 1939

CONTENTS

I. INTRODUCTION

In considering the results of x-ray investigations of crystals one should distinguish clearly between two cases: (*1*) If the investigated substance is simple and if large well-built crystals of high symmetry are obtainable, the analysis can be carried out using the most modern x-ray methods, interpreted with the aid of structure interference theory. In such a case the diagrams lead, without additional help, to the exact positions of all atoms in the crystal lattice. In the early days of x-ray investigation of crystal structures, W. H. and W. L. Bragg (4) determined the structures of diamond, rock salt, and calcite in this way. More recently, comparatively complicated structures have also been completely worked out. Usually the results of such a thorough investigation are presented in the form of a map showing the electron density throughout the elementary cell. Robertson (43), to whom we owe many of these complete structure determinations, has recently given a survey on this development.

(*2*) Up to the present time, however, only a comparatively small number of crystal structures have been successfully worked out in this way, using solely data from the x-ray measurements. Frequently the low symmetry of the crystals prevents such a straightforward analysis. In other cases

[1] Presented at the Symposium on X-ray Studies of Substances of High Molecular Weight, held at the Ninety-eighth Meeting of the American Chemical Society in Boston, Massachusetts, September, 1939, under the auspices of the Divisions of Physical and Inorganic Chemistry, Colloid Chemistry, Organic Chemistry, Paint and Varnish Chemistry, and Rubber Chemistry of the American Chemical Society. Because of the unavoidable absence of Dr. Mark, the paper was read by Dr. Charles W. Stillwell.

single crystals of sufficient quality cannot be prepared, or there are so many atoms in the elementary cell that the mathematical difficulty of a direct analysis becomes insurmountable. It is then necessary to carry out the investigation with the aid of other physical or chemical properties of the substance being studied.

First of all, the chemical formula itself, especially in the case of organic materials, may offer certain reliable starting points for the development of a structural picture. Sometimes the optical behavior of the substance—its double refraction and rotatory power—may give additional clues for the investigation. In special cases the hardness and the tensile strength of the investigated material have given valuable additional information.

In such indirect determinations of structure it has proved to be profitable not to confine the investigation to a single substance but to extend it to a whole group of compounds which are structurally closely related. This procedure has been very successfully used by A. Mueller and coworkers (35)[2] in studying long-chain organic compounds and by W. L. Bragg (5)[3] and his collaborators in their extensive studies of the structure of silicates.

In the case of high polymeric substances—especially cellulose—one must also use this second method. It is necessary to try to obtain as much reliable x-ray diffraction information as possible, not only from cellulose in its different native states but also from its derivatives, degradation products, and even from glucose and cellobiose, the fundamental units of cellulose. In addition, it has proved to be important to get as much x-ray evidence as possible on the structures of the lower carbohydrates,—those belonging to the group of tetroses, pentoses, and hexoses. These studies of sugars of low molecular weight have given a sound fundamental basis for the inferences drawn in the field of high polymeric carbohydrates.

The interpretation of the x-ray diagrams has been successfully accomplished only by taking into account all the chemical evidence concerning the structure of organic substances, especially the laws of stereochemistry and the absolute dimensions of the atoms in Ångström units.

This paper gives a short survey of our present knowledge of the structure of all carbohydrates up to native and mercerized cellulose, while the following paper by Dr. Sisson shows how this knowledge can be evaluated in the analysis of orientation and swelling of cellulose samples and how the structure of this material is connected with its technical and biological properties.

[2] For a comprehensive survey see reference 54, pages 135–95.
[3] For a comprehensive survey see reference 54, pages 106–29.

II. INVESTIGATIONS IN THE ERYTHRITOL GROUP

The simplest substance which may be regarded as belonging to the large group of sugars and which has been investigated by x-rays is *inactive* erythritol, $CH_2OH \cdot CHOH \cdot CHOH \cdot CH_2OH$. The investigation showed that the molecule has the form of an open chain with four carbon atoms and that like substituents attached to the middle carbon atoms are *trans* with respect to each other (7, 48).

The symmetry of the single molecule inside the crystal lattice is that of a center of symmetry (C_i). This is interesting, owing to the fact that we have an inactive compound before us which contains two asymmetric carbon atoms. These atoms have the same rotatory power but with inverse sense and hence compensate each other. An inner molecular compensation of the rotatory power can be effected by a plane of symmetry or by a center of symmetry. In the first case it would mean that the molecule itself shows the *cis*-configuration, while the latter indicates that the arrangement inside the molecule points to *trans*.

It was first pointed out by Reis that a center of symmetry is much more probable if one assumes that polar forces of any kind determine the spatial arrangement of the different parts of a large molecule, because a center of symmetry brings groups of equal charge further apart and brings groups of opposite charge nearer to each other. This is just what one would expect to obtain if equal charges repel each other and opposite charges attract.

As far as present evidence goes, large molecules with inner molecular compensation of rotatory power always have a symmetry center; a plane of symmetry has not yet been observed.

Much work has been done on pentaerythritol and its derivatives (15, 21, 26, 28, 36, 38, 52). The first investigations led to the conclusion that the symmetry of this molecule inside the crystal lattice was pyramidal. This would be in definite disagreement with the expectations of stereochemistry, from which one would predict a tetrahedral arrangement of the four equivalent substituents. A very careful examination of the situation, made by many authors, finally showed that the first x-ray investigation was correct but that the crystal class cited in the literature was wrong. A redetermination of the crystal class gave S_4. Together with the x-ray data this leads to a tetrahedral arrangement of the four groups.

To make this quite sure, a number of derivatives, such as pentaerythritol tetrachloride, pentaerythritol tetrabromide, pentaerythritol tetraiodide, pentaerythritol tetranitrate, pentaerythritol tetraacetate, and pentaerythritol tetraformate, have been investigated (6, 11, 24, 51). They crystallize in the tetragonal and rhombic systems. In all cases a complete symmetry analysis of the molecules could be carried out, which showed that all of these molecules show a symmetry in complete agreement with

the expectations of stereochemistry. So far, this group of sugars fits exactly into the framework of classical stereochemistry, without providing any surprises.

Another derivative of pentaerythritol which crystallizes in the hexagonal system is dibenzalpentaerythritol. The symmetry analysis of this substance showed that the three valence directions of the carbon atom marked by a cross in figure 1 are in one plane (30). This means a rather strong deformation of the original valence directions, which may be attributed to the presence of the two rings on each side of the central carbon atom. We have before us a molecule of the spirane type, where valence distortions are not unfamiliar.

All these investigations lead to the conclusion that the atomic diameters characteristic of the carbon and oxygen atoms are strictly constant in carbohydrate structures, though there may be valence angle distortions producing considerable deviations from the normal tetrahedral configuration.

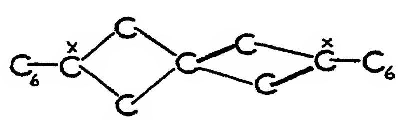


FIG. 1. Structure of dibenzalpentaerythritol

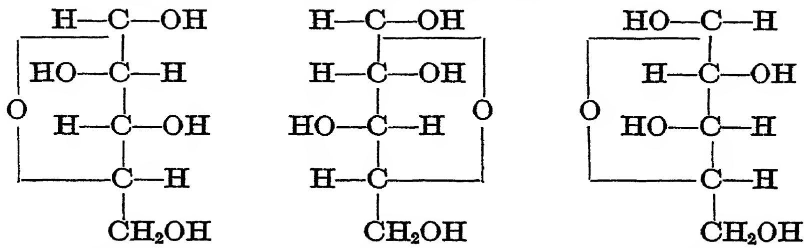
III. SUGARS WITH FIVE, SIX, AND MORE CARBON ATOMS

Molecules of this type may be regarded as especially interesting and important, owing to the fact that *β-d*-glucose, the elementary chemical unit of cellulose, belongs to this group. Two possible arrangements have to be kept in mind for these molecules,—the open-chain form and the closed-ring form. Both have been found during the investigation of well-crystallized low molecular substances, in which the most advanced x-ray technic has been utilized.

The two isomeric alcohols dulcitol and mannitol, having the formula $C_6H_{14}O_6$, have been carefully investigated. Dulcitol crystals are monoclinic; those of mannitol are rhombic. In both cases the elementary cell contains four asymmetric molecules. From intensity measurements it has been shown that each of these molecules has the form of a straight zigzag chain, the diameter of the carbon atom being 1.5 Å. This is in complete agreement with the conclusions reached from the investigation of long-chain paraffins and their derivatives. It shows that the presence of the hydroxyl groups does not seriously affect the shape of the carbon skeleton (29).

Another set of isomeric compounds having five carbon atoms which has

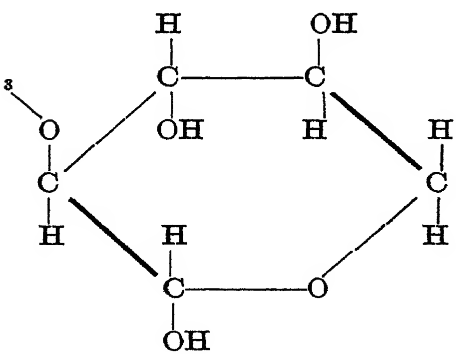
been thoroughly investigated includes *α-l*-xylose, *α-d*-xylose, and *β-l*-arabinose. These have the following formulas:



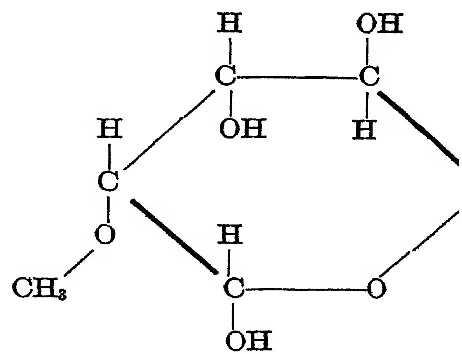
α-l-xylose *α-d*-xylose *β-l*-arabinose

All three form rhombic crystals with four molecules in the elementary cell. The molecules are asymmetric. From the intensity distribution of the scattered radiation it has been deduced that in each case the molecule is a ring with five links, having approximately the following dimensions: 6.5, 5.6, and 4.8 Å. This shows that we have to regard such sugar molecules as being elliptic plates, the thickness of which is slightly less (8, 30) than the smaller diameter. As will be mentioned later, this conception agrees completely with the chemical evidence on the structures of simple sugars (2, 8).

In addition to the xyloses, the *α*- and *β*-methylxylosides, having the formulas



and



have been studied. Both form perfect monoclinic crystals, well-suited to x-ray investigation. In each case the molecule has the form of a ring with six links, containing five carbon atoms and one oxygen atom. The five carbon atoms lie very closely in one plane, while the oxygen atom which closes the ring lies about 0.8 Å. above this plane. This shows again that in building up structures of sugars we may rely firmly on the general rules concerning the diameters of carbon and oxygen atoms, but we have to

consider certain deviations of the valence directions, a fact which will be discussed more in detail later.

After these preliminary remarks concerning sugars of low molecular weight, we turn now to the fundamental unit of cellulose, namely glucose, and its derivatives.

β-d-Glucose crystallizes in the rhombic system. The elementary cell contains four molecules. Consideration of the x-ray diagrams shows that the molecule has the shape of a ring with six links, being built up of five carbon atoms and one oxygen atom. No evidence regarding details of the ring is available (16).

An interesting investigation was carried out by Nowakowsky (39) with three esters of glucose: namely, *α*-acetylglucose, *α*-laurylglucose, and *α*-palmitylglucose. These compounds crystallize in needles which, bundled together, give more or less distinct fiber diagrams. This fact made it possible to determine the fiber period; it was found to be practically the

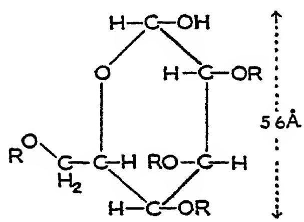


FIG. 2. Formula of the tetra esters of glucose

same in all three cases, being 5.39, 5.36, and 5.38 Å., respectively. On the other hand, the edge lengths of the unit cells perpendicular to the fiber axis were entirely different and increased from the acetate over the laurate to the palmitate. This shows that in these structures the glucose ring runs parallel to the needle axis, as shown in figure 2, and that the substituent is extended perpendicular to this direction. The period of about 5.4 Å. coincides with the length of one diameter of the xylose ring mentioned above. It seems that the ring diameter between carbon atoms 1 and 4 corresponds to this distance. It will later be shown that the distance between the center of gravity of these two atoms is only about 3 Å. but the substituents allotted to them increase the dimensions of the ring in this direction to about 5.4 to 5.6 Å.

The fact that the substituting groups lie perpendicular to the glucose ring is significant and has been confirmed in the investigation of the different esters of cellulose.

One step closer to cellulose is *β-d*-cellobiose, the structure of which is shown in figure 3. This substance crystallizes in the monoclinic system.

The elementary cell contains two molecules. Owing to the existence of small but well-formed crystals, a comparatively careful investigation has been possible. It points to the fact that the cellobiose molecule is composed of two rings which are practically coplanar. They are held together by an oxygen atom which is about 0.8 Å. from the plane of the rings. This structure is closely related to that of the cellobiose residue in cellulose chains (16).

The transition from the molecular lattice of cellobiose to the main-valence chain lattice of cellulose is of great interest, and its careful experimental investigation would be very welcome. Unfortunately, cellobiose

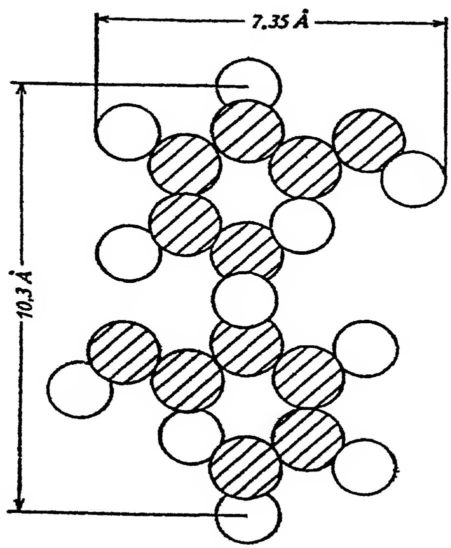


FIG. 3. Structure of β-*d*-cellobiose; hydrogen atoms are omitted

and its derivatives are the last sugars in the cellulose family which can be obtained in single crystals. The higher representatives of the cellodextrins are available only as crystal powders, a fact which prevents their complete x-ray investigation. Nevertheless, considerable work has been concentrated on them (9, 50, 53). First of all, it may be emphasized that the investigations of Willstätter, Zechmeister, and Toth have shown without a doubt that three well-defined sugars can be prepared: namely, cellotriose, cellotetrose, and cellohexose. They were separated by hydrolyzing cotton and fractionating the product yielded thereby. These sugars can be methylated and acetylated without destroying their structure.

All chemical and physical investigations have led to the conviction that

they are homogeneous, pure substances with three, four, and six glucose rings in the chain.

X-ray diagrams of the sugars themselves have furnished only powder diagrams which could not be interpreted. An investigation of the acetates, however, has yielded some results. From an x-ray study of glucose pentaacetate, cellobiose octaacetate, and cellotriose undecaacetate, the dimensions of the three elementary cells of these substances have been deduced (25, 55). The figures are given in table 1. They show that the first axis of the elementary cell is practically the same in all these three compounds. This distance must apparently be perpendicular to the chain direction and hence perpendicular to the glucose ring, as shown in

TABLE 1

Dimensions of the elementary cells of three acetates

SUBSTANCE	NUMBER OF MOLECULES IN THE ELEMENTARY CELL	IDENTITY PERIODS OF THE ELEMENTARY CELL		
		a	*b*	*c*
		Å.	Å.	Å.
Glucose pentaacetate...	4	5.65	14.9	24
Cellobiose octaacetate..........	4	5.7	15	42
Cellotriose undecaacetate.. ...	4	5.7	15	60

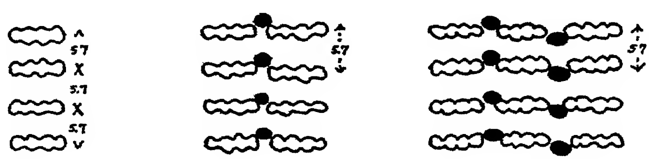


FIG. 4. Comparison of the structures of the acetates of glucose, cellobiose, and cellotriose

figure 4. The second axis of the elementary cell is also practically constant; it is apparently related to one ring diameter perpendicular to the chain of cellobiose and cellotriose. This ring diameter would normally be equal to a distance of about 6.3 Å. Here, however, it is more than double this value, owing to the fact that all hydroxyl groups of the three sugars have been substituted with acetic acid. The third axis of the elementary cell shows an increase from glucose to cellobiose to cellotriose, the increment being about 18 Å. Apparently this is the direction in which the chain is developing, as roughly sketched in figure 4.

These three acetates crystallize in needles. From a bundle of the needles one obtains comparatively good fiber diagrams, from which the identity period along the needle axis is calculated to be 5.72 Å. in the case of glucose

and 5.59 Å. in the case of cellobiose. These distances coincide, within experimental error, with the identity periods in the a direction (table 1), indicating that the needle axis is perpendicular to the direction in which the chains develop and perpendicular to the plane of the glucose ring. This is in good agreement with our knowledge of the structure of other organic crystals, e.g., fatty acids, naphthalene, anthracene, etc., showing again that as long as normal molecular lattices are prevailing, the direction of the largest development of the crystals (in this case the needle axis) is perpendicular to the direction of the largest development of the molecule (in this case the chain axis). The same conclusion has been reached in studies of fatty acids; in these compounds the molecular chains lie perpendicular to the plane of the flat crystal sheets.

In compounds containing a main-valence lattice, on the other hand, the chain axes of the molecules are parallel to the fiber axis of the sample.

Although there is no visible similarity between the diagram of cellulose on one hand and that of cellobiose and glucose on the other, the investiga-

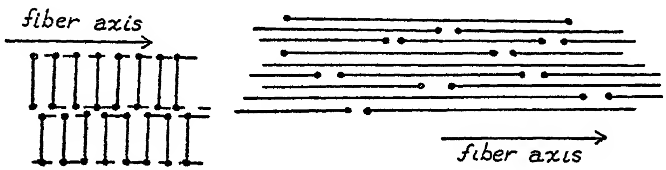


FIG. 5. a, molecular lattice, with chains of identical length; b, chain lattice, with molecules of different length

tion of a mixture of higher cellodextrins which had been produced by hydrolyzing cotton showed their x-ray diagrams to be completely analogous to those of mercerized cellulose. It is easy to see that, if one has long chains bundled together into a lattice-like arrangement, the ends of the chains will be distributed irregularly all over the crystallized areas and hence will not contribute anything to the interference phenomenon. Whether such a structure is composed of long or short chains will not affect the character of the x-ray reflections in any respect. However, as soon as all the chains are of identical length a normal molecular lattice is built up and the chain ends are all concentrated in certain lattice planes. As a consequence, new x-ray reflections appear and a different diagram is observed. (Figure 5 shows diagrammatically the difference between a substance with chains of equal length forming a normal molecular lattice, in which the chains are perpendicular to the fiber axis, and a material which has a main-valence chain lattice, in which the chains run parallel to the fiber.)

Summarizing we may say: All x-ray investigations on well crystallizing

carbohydrates show that in the lattices of these substances the atomic distance rules are maintained, although there may be considerable distortions of the directions of the valence forces. All carbohydrates and their derivatives which have been investigated have normal molecular lattices. The lower sugars, including glucose and cellobiose, have ring structures. The dimensions of the rings have been determined. The x-ray study of cellotriose, cellotetrose, cellohexose, and some other dextrins has led to a reasonable picture for the transition from a normal molecular lattice to a structure built up of chains of irregular length which are held together by intermolecular forces.

IV. X-RAY INVESTIGATIONS OF CELLULOSE

The fact that cellulose gives a definite x-ray diagram was first detected in 1913 by Nishikawa and Ono (37). In 1918 Scherrer (46) investigated cellulose with x-rays and reported a fiber diagram in 1920 (47). It was clear that at that time no attempt to evaluate these diagrams could be successful, because our general knowledge of crystal structure and x-ray scattering was much too incomplete. As a matter of fact, Nishikawa, Ono, and Scherrer confined themselves to the statement that natural fibers show a very interesting x-ray scattering effect, indicating the presence of small oriented crystals.

Beginning in 1919, this problem was attacked experimentally by Herzog and Jancke (17). These workers confirmed Scherrer's first statement and started a series of systematic investigations with different kinds of cellulose fibers. They discovered a new type of x-ray diagram, consisting of a group of symmetrically distributed points or stripes. This they called a "fiber diagram." It was first interpreted theoretically by Polanyi (40), who was later aided in his work by Weissenberg (42). In this way an important experimental and theoretical advance was made, which encouraged the drawing of some conclusions concerning the structure of the investigated material from the new experimental evidence. Without chemical help reliable deductions were clearly impossible; the assumption was therefore made that cellulose is built up from glucose residues and that cellobiose exists to a certain extent in the lattice.

These considerations did not lead to entirely definite results, but offered three different possibilities for the structure of cellulose. It is very interesting and certainly should be clearly emphasized that one of the structures proposed by Herzog (18) and Polanyi (41) was a continuous chain of glucose residues very similar to the solution which is regarded today as being the best. It is quite evident that at that early date (1921) it was impossible to choose definitely between the different models deduced from the x-ray studies without the aid of arguments coming from another type of investigation.

It is interesting to note that in the same year, 1921, such an argument was in fact brought forward by Freudenberg as a result of chemical experiments with methylcellulose (10). He came to the conclusion that, according to the results of methylation and degradation, cellulose should be composed of long chain-like molecules.

In this way both organochemical and physicochemical measurements pointed in the same direction but no successful combination of these results was attempted at that time. This may be explained by the fact that the deductions from the x-ray measurements were published very cautiously and tentatively, without emphasizing the possibility of long chain molecules and by the fact that at the same time other chemical investigations seemed to point more in the direction of small ring-shaped units in the cellulose lattice.

It was necessary that the chemical situation undergo a more complete clarification and that our general knowledge regarding the structure of organic crystals be increased. Progress was soon made in both these directions. Of special importance was Haworth's definite establishment of a structural formula for the glucose residue. It was shown to be a ring built up from five carbon atoms and one oxygen atom (13, 14). At the same time cellobiose was shown to be built up of two glucose units linked together by a 1,4-glucosidic main-valence bond.

As a result of many x-ray investigations of other organic substances, the atomic diameters of the carbon and oxygen atoms could be determined. In particular, W. H. and W. L. Bragg proposed that in all organic lattices the carbon atom has a diameter of about 1.5 Å. and the oxygen atom of about 1.4 Å. If one combines these values with the ideas of stereochemistry it is possible to build up molecular models which not only show diagrammatically the shape of the particle but which give a true picture of the dimensions and the form of a given organic molecule. This was successfully carried out in several cases, as with naphthalene, anthracene, fatty acids, etc.

It was obvious that application of the recently acquired knowledge of sugar chemistry (Haworth), together with the new information about molecular dimensions (W. L. Bragg), to the cellulose problem as left in 1921 (Freudenberg, Herzog, Polanyi) would have a good chance of leading to a successful solution of the problem of cellulose structure.

Such an attempt was, in fact, made by Sponsler and Dore in 1926 (49). These authors gave the first spatial molecular model of cellulose based upon the possibility of infinitely long glucose chains. Unfortunately, they selected a type of linkage which did not agree with the chemical evidence. However, as shown by Meyer and Mark (32), the original model of Sponsler and Dore could easily be changed so as to make it satisfy all physical and chemical requirements.

This finished the first step in the development of a suitable structural formula for cellulose and established the fundamental principle of its structure. The crystallized regions are composed of long main-valence chains of glucose residues linked together by 1,4-glucosidic bonds. This statement has never met any experimental opposition but, on the contrary, may be regarded as being supported by all the evidence available up to the present time.

After this achievement a second period of cellulose structure work in the x-ray field started. It was evident that the models proposed by Polanyi, Sponsler and Dore, and Meyer and Mark had to be regarded as first approximations. The next task was to work out as accurately as possible the positions of all the atoms in the elementary cell of native cellulose.

This elementary cell had first been assumed to be rhombic (Polanyi; Sponsler and Dore). Later a monoclinic cell was proposed (Meyer and Mark), and another monoclinic cell has been considered by Sauter (45). From the experimental point of view, the most important factor is the use of highly oriented preparations which not only show fiber structure but which are also oriented in respect to the other two crystallographic axes. Using the most highly oriented material available and the technique most appropriate for the purpose, Gross and Clark (12) have obtained what may be regarded as a final solution of this important question. They conclude that a monoclinic cell with the dimensions a = 8.35 Å., b = 10.3 Å., c = 7.95 Å., and β = 84° is in best agreement with all experimental evidence to date.

The space group of cellulose can also be regarded as definitely settled; it is C_2^2.

In this group there are two sets of twofold screw axes parallel to the b-axis. These are independent of each other. The two parallel sets of glucose chains which run through the elementary cell can therefore be disposed in two different ways: (a) they may be oriented in the same direction, or (b) they may be oriented in opposite directions. To choose between these possibilities, a very careful study of the intensity distribution in the x-ray diagrams is necessary. This was made by Mark and v. Susich (27), and their work was extended by Andress (1). These authors showed that the assumption that all the chains are oriented similarly was in satisfactory agreement with the experimental evidence then available. Later, new experiments showed that there was more probability for arrangement b, with the two sets of main-valence chains running in opposite directions. In conformity with this conclusion, Meyer and Misch (34) proposed for cellulose the model represented in figures 6, 7, 8, and 9.

It seems to be reasonable to discuss this model briefly and to consider to

what extent it meets our present requirements from chemical and physical points of view.

(*a*) *Optical behavior.* This model explains the double refraction and the rotatory power of cellulose in a qualitative way, owing to the fact that its anisotropy is that of a uniaxial crystal.

(*b*) *Mechanical behavior.* The fact that the main-valence chains are parallel to the fiber axis explains the high tensile strength in this direction and the smaller strength perpendicular to it. The modulus of elasticity of cellulose, as determined by Meyer and Lotmar (31), is in good agreement with this model. The thermal expansion, which has been measured by Hengstenberg, agrees equally well (33).

(*c*) *X-ray behavior.* All observed x-ray interference spots can be completely explained, in respect to their position, intensity, and width. No exception has yet been found.

(*d*) *Swelling.* The anisotropy in swelling confirms the main-valence chain model of cellulose and has recently led to important progress concerning the process of mechanical deformation.

The question of intramicellar swelling leads to the problem of the nature of the forces by which the lattice of the cellulose is kept together. It is well known that cellulose (as well as other high molecular compounds) has the peculiar quality of reacting with different chemical substances without losing its lattice structure. This topochemical reactivity must have some connection with the lattice structure and the forces which keep this structure together.

From an examination of figures 6 to 9 we can derive the following conclusions: In the direction of the *b*-axis the main-valence chains are kept together by 1,4-glucosidic bonds, the strength of which may be regarded as being of the order of magnitude of 50,000 calories per mole (see especially figures 8 and 9).

Considering the forces which hold the structure together in the other two directions we first concentrate our attention on figure 7. This shows that along the *a*-axis the glucose rings are in rather close proximity to each other. As a matter of fact, the centers of the two nearest oxygen atoms of two such groups approach each other to about 2.5 Å. It is well known from our general knowledge of organic crystal structures that two oxygen atoms belonging to different molecules should not approach nearer than 3.0 Å. if there are only van der Waals' forces acting between them. The two oxygen atoms of the glucose residues in the cellulose lattice are much closer. This arouses the suspicion that stronger forces are acting between them in the lattice of cellulose. It seems to be reasonable to assume that a hydrogen bond is established between these two oxygen atoms. According to Huggins (20) this would correspond to a strength of about 15,000 calories

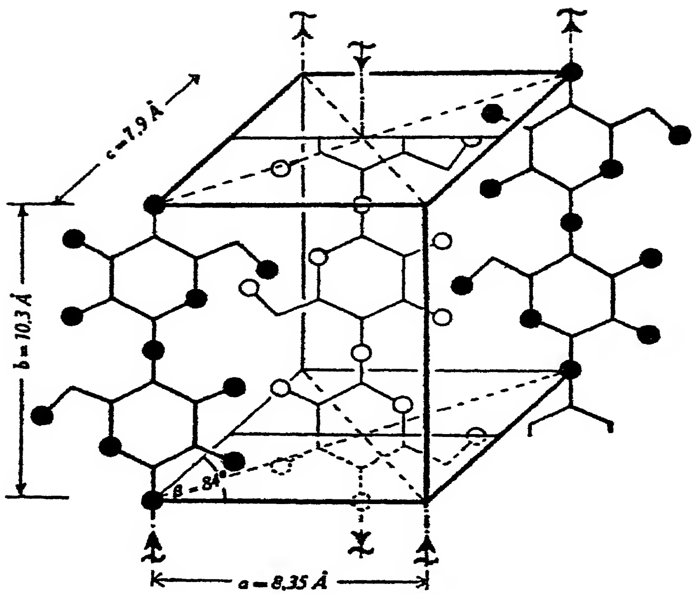


FIG. 6. Representation of the monoclinic elementary cell of cellulose (according to Meyer and Misch). Three chains of glucose units are shown, running parallel to the *b*-axis (fiber axis).

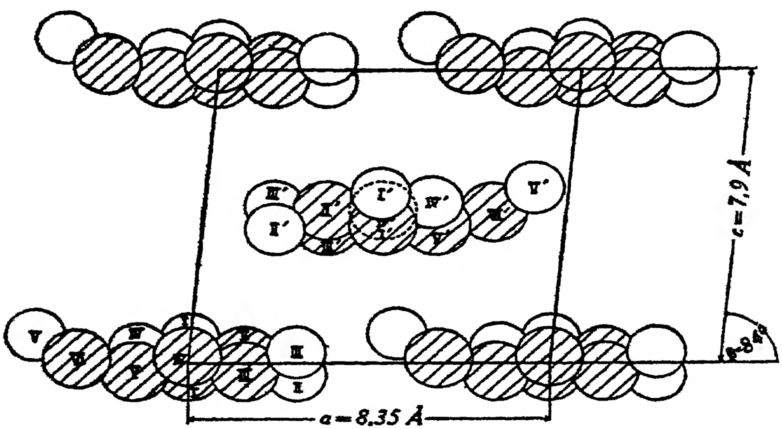


FIG. 7. View of the elementary cell of cellulose along the main-valence chains (*b*-axis). One sees the projections of the glucose groups on the *a*-*c* plane. Hydrogen atoms are omitted.

per mole. This would explain the comparatively tight packing in the *a*-*b* plane in the direction of the *a*-axis and would, at the same time, offer an explanation for the fact that more highly oriented structures can be

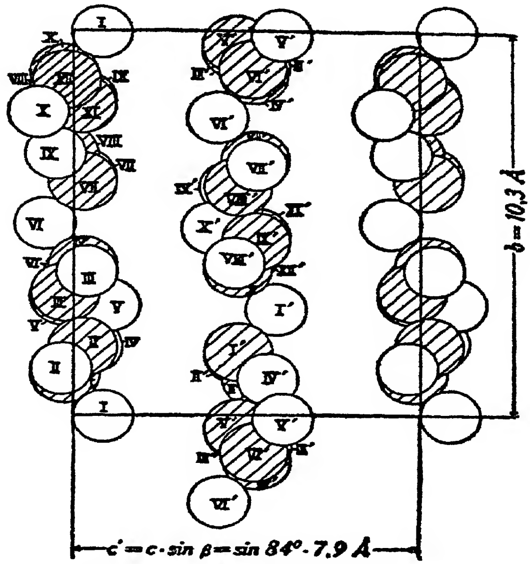


FIG. 8. View of the elementary cell of cellulose along the *a*-axis. One sees the chains from the side and observes that rather large spaces are between them. Hydrogen atoms are omitted.

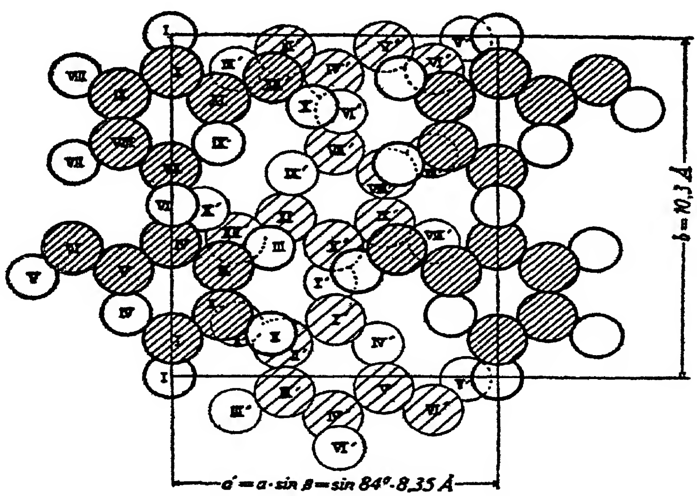


FIG. 9. View of the elementary cell of cellulose along the *c*-axis. One looks at the main-valence chains from their broad side. Hydrogen atoms are omitted.

obtained by mechanical treatment. Such an arrangement would be, at the same time, in agreement with the expectations of Bernal (3), starting from the discussion of the structure of ice.

We finally ask what forces hold together the lattice along the *c*-axis. According to figures 7 and 8, the nearest distance of atomic centers in this direction is about 3.1 Å. (This is the distance between a carbon atom of one chain and an oxygen atom of the other. The nearest distance between hydroxyls is 3.8 Å.) This corresponds closely to the distance to be expected if van der Waals' forces hold the lattice together in this direction. Such forces correspond to an energy of about 8000 calories per mole.

From this point of view, the lattice of cellulose may be regarded as a combination of a *chain lattice* and a *layer lattice*. The strongest forces act along the *b*-axis, but the forces along the *a*-axis are also comparatively large. There are main-valence chains along the *b*-axis and hydrogen bond nets in the *a-b* plane. Perpendicular to this plane, however, weak forces are acting and the spacing between these planes is large. This permits chemical reagents to penetrate easily into the lattice and to react with the hydroxyl groups of cellulose.

It would be premature to propose further and more detailed consequences of this viewpoint, but it might not be too far-reaching to say that the fact that cellulose is held together by three different kinds of forces in three different directions of space is responsible for the interesting and important properties which this substance exhibits during its reactions in the solid state.

V. X-RAY INVESTIGATIONS OF STARCH

Although no conclusive quantitative results could be derived from x-ray measurements on starch, this short report on x-ray investigations of carbohydrates would be incomplete if a few words and some literature references were not added concerning the structure of starch.

Starch has frequently been investigated with x-rays, but the results have not been very encouraging. One always obtains either a diagram which can only be called an amorphous halo or a few interference rings showing that there is some crystalline or micellar structure prevailing in the investigated sample.

However no attempt to evaluate these diagrams quantitatively has been successful hitherto, the main reason being that in spite of all efforts no oriented sample could be obtained which would allow a reliable calculation of the elementary cell. One is confined to measuring and evaluating a few diffused interference rings without having any other indication regarding the symmetry and dimensions of the lattice.

In the last few years systematic experiments have been carried out by Katz (22, 23, 44) concerning the changes which the x-ray diagram of starch undergoes if the sample is treated in different ways. He distinguishes between a crystallized and an amorphous diagram and shows that they can be converted into each other under certain conditions.

The only result which can be derived from the x-ray investigation of starch is a negative one. The impossibility of getting fiber diagrams, even after extreme mechanical treatment of the samples, indicates that apparently there are no long unbranched chain molecules present in starch. In this respect the x-ray investigations point in the same direction which has been furnished with very much stronger arguments in the last few years by chemical considerations.

The methylation and subsequent degradation of starch has led Hirst (19) to the conclusion that particles of starch consist of a highly branched chain system, in which the average individual length of the branches corresponds to a molecular weight of about 4500. X-ray measurements to date cannot supply any quantitative additional knowledge to this picture, but it can be stated that they are not contrary to it.

REFERENCES

(1) Andress, R. K.: Z. physik. Chem. **B2**, 380 (1929).
(2) Andress, R. K., and Reinhardt, L.: Z. Krist. **78**, 477 (1931).
(3) Bernal, J. D.: Proc. Roy. Soc. (London) **A151**, 384 (1936).
(4) Bragg, W. H., and Bragg, W. L.: Proc. Roy. Soc. (London) **A89**, 277 (1913).
(5) Bragg, W. H., and Warren, B. E.: Z. Krist. **69**, 168 (1928); **76**, 201 (1930).
(6) Bredig, M. A.: Z. Krist. **74**, 49 (1930).
(7) Burgers, W. G.: Phil. Mag. **1**, 289 (1926).
(8) Cox, E. G.: J. Chem. Soc. **1931**, 2313.
(9) Dziengel, K., Trogus, C., and Hess, K.: Ber. **65**, 1454 (1932).
(10) Freudenberg, K.: Ber. **54**, 770 (1921).
(11) Gerstäcker, A., Moeller, H., and Reis, A.: Z. Krist. **66**, 355 (1927).
(12) Gross, S. T., and Clark, G. L.: Z. Krist. **A99**, 357 (1938).
(13) Haworth, W. N.: Nature **116**, 430 (1925); J. Chem. Soc. **128**, 89 (1926).
(14) Haworth, W. N., and Machemer, H.: Ber. **65**, 60 (1932).
(15) Hendricks, S. B.: Z. Krist. **66**, 131 (1927).
(16) Hengstenberg, J., and Mark, H.: Z. Krist. **72**, 301 (1929).
(17) Herzog, R. O., and Jancke, W.: Z. Physik **3**, 196 (1920); Ber. **53**, 2162 (1920).
(18) Herzog, R. O., and Jancke, W.: Z. angew. Chem. **34**, 213, 385 (1921).
(19) Hirst, E. L., and Young, G. T.: J. Chem. Soc. **1939**, 1471.
(20) Huggins, M. L.: J. Org. Chem. **1**, 407 (1936).
(21) Huggins, M. L., and Hendricks, S. B.: J. Am. Chem. Soc. **48**, 164 (1926).
(22) Katz, I. R., and Derksen, I. C.: Z. physik. Chem. **A166**, 27 (1933); **A168**, 321 (1934); **A175**, 383 (1936); **A180**, 405 (1937).
(23) Katz, I. R., and Weidinger, A.: Z. physik. Chem. **A180**, 423 (1937).
(24) Knaggs, I. E.: Mineralog. Mag. **20**, 346 (1925); Proc. Roy. Soc. (London) **A122**, 69 (1929).
(25) Leuck, G. J., and Mark, H.: J. Am. Chem. Soc. **56**, 1959 (1934).
(26) Mark, H., and v. Susich, G.: Z. Krist. **69**, 105 (1928).
(27) Mark, H., and v. Susich, G.: Z. physik. Chem. **B4**, 431 (1929).
(28) Mark, H., and Weissenberg, K.: Z. Physik **17**, 301 (1923).
(29) Marwick, T. C.: Proc. Roy. Soc. (London) **A131**, 621 (1931).
(30) v. Melle, F. A., and Schurink, H. B. J.: Z. Krist. **69**, 1 (1928).
(31) Meyer, K. H., and Lotmar, W.: Helv. Chim. Acta **19**, 65 (1936).

(32) Meyer, K. H., and Mark, H.: Ber. **61**, 593 (1928); Z. physik. Chem. **B2**, 115 (1929).
(33) *cf.* Meyer, K. H., and Mark, H.: *Der Aufbau der hochpolymeren organischen Naturstoffe*, p. 97. Akademische Verlagsgesellschaft, Leipzig (1930).
(34) Meyer, K. H., and Misch, L.: Helv. Chim. Acta **20**, 232 (1937).
(35) Mueller, A.: Proc. Roy. Soc. (London) **114**, 542 (1927); **120**, 437 (1928).
(36) Nehmitz, A.: Z. Krist. **66**, 408 (1927).
(37) Nishikawa, S., and Ono: Proc. Phys. Math. Soc. Japan, September 20, 1913.
(38) Nitta, G.: Bull. Chem. Soc. Japan **1**, 62 (1926).
(39) Nowakowsky, A.: Compt. rend. **191**, 411 (1930).
(40) Polanyi, M.: Z. Physik **7**, 149 (1921).
(41) Polanyi, M.: Naturwissenschaften **9**, 288 (1921).
(42) Polanyi, M., and Weissenberg, K.: Z. Physik **9**, 123 (1922).
(43) Robertson, J. M.: Reports on Progress in Physics **4**, 332 (1938).
(44) Samec, M., and Katz, I. R.: Z. physik. Chem. **A180**, 436 (1937).
(45) Sauter, E.: Z. physik. Chem. **B35**, 83 (1937).
(46) Scherrer, P.: Göttinger Nachrichten 1918, p. 98.
(47) Scherrer, P.: In R. Zsigmondy's *Kolloidchemie*, 3rd edition. O. Spamer, Leipzig (1920).
(48) Schonfeld, N., Hermann, K., and Hassel, O.: Z. physik. Chem. **A124**, 305 (1926).
(49) Sponsler, L. O., and Dore, W. H.: Colloid Symposium Monograph **4**, 174 (1926).
(50) Toth, G.: Papier Ztg. **56**, 1 (1938) (a comprehensive paper).
(51) Wagner, G., and Dengel, G.: Z. physik. Chem. **B16**, 382 (1932).
(52) Westenbrink, H. G. K., and v. Melle, F. A.: Z. Krist. **64**, 548 (1926).
(53) Willstätter, R., and Zechmeister, L.: Ber. **62**, 722 (1929).
(54) Wyckoff, R. W. G.: *The Structure of Crystals*, Supplement to the 2nd edition. Reinhold Publishing Corporation, New York (1935).
(55) Zechmeister, L., Mark, H., and Toth, G.: Ber. **66**, 269 (1933).

X-RAY STUDIES REGARDING THE STRUCTURE AND BEHAVIOR OF NATIVE CELLULOSE MEMBRANES[1]

WAYNE A. SISSON

Cellulose Department, Chemical Foundation, Boyce Thompson Institute for Plant Research, Inc., Yonkers, New York

Received December 16, 1939

CONTENTS

---

[1] Presented at the Symposium on X-ray Studies of Substances of High Molecular Weight, held at the Ninety-eighth Meeting of the American Chemical Society in Boston, Massachusetts, September, 1939, under the auspices of the Divisions of Physical and Inorganic Chemistry, Colloid Chemistry, Organic Chemistry, Paint and Varnish Chemistry, and Rubber Chemistry of the American Chemical Society.

I. INTRODUCTION

X-ray diffraction analysis, as a research tool, has two closely related fields of application. The first is concerned with the form, properties, and three-dimensional arrangement of atoms and molecules within a single crystal; the second is concerned with the form, arrangement, and properties of crystalline aggregates in a polycrystalline material.

Since the x-ray diffraction diagram of a native cellulose fiber consists principally of definite diffraction rings, one is forced to conclude that the major portion of the fiber consists of crystalline cellulose. The first field of x-ray analysis, therefore, as applied to cellulosic materials, is the interpretation of the stereochemistry of the cellulose molecule within the crystal lattice. This structure, which may be referred to as "cellulose structure," has been discussed by Mark (29) in the preceding paper of this symposium.

From the fact that cellulose fibers give continuous diffraction rings (i.e., Debye–Scherrer diagrams), a second important deduction may be made: namely, that the crystallinity of the fiber is discontinuous. In other words, a cellulose fiber is not a single crystal; it is a crystalline aggregate. This is definitely proven by the fact that even a single fiber, when photographed by a micro method (7, 37), gives smooth rings instead of Laue spots. These small crystals of cellulose which build up the cell wall of a fiber will be referred to as "cellulose crystallites," and the material between these crystalline areas as "intercrystalline material." The second field of x-ray analysis, therefore, as applied to cellulosic materials, is the interpretation of the form, arrangement, and properties of the cellulose crystallite and its relation to the intercrystalline material within the plant cell wall. This dual structure built up of crystalline and intercrystalline material, which is discussed in the present paper, will be referred to as "membrane structure."

The x-ray data used in the interpretation of cellulose membrane structure are of two types: (*1*) certain characteristics of the x-ray diagram of the untreated fiber, and (*2*) the changes produced in these characteristics by the action of various physical and chemical treatments of the fiber.

II. CHARACTERISTICS OF CELLULOSE DIFFRACTION PATTERNS

X-ray diagrams may be described and differentiated one from the other in terms of the following characteristics: (A) Number of diffraction rings; (B) diameter of each ring; (C) relative intensity of each ring; (D) concentration of rings into arcs,—orientation; (E) width or diffuseness of each ring; and (F) general scattering,—amorphous pattern.

The number (A), diameter (B), and relative intensity (C) of the diffrac-

tion rings appear to be constant for all native cellulose fibers. Since these characteristics are associated with crystalline structure, it may be concluded that the crystalline structure of cellulose is the same for each fiber. The arcing (D) and width (E) of each diffraction ring, and the general scattering (F), however, are not the same for each fiber, and since these characteristics are associated in part with the membrane structure, it may be concluded that this structure varies from fiber to fiber. Any concept of cellulose membrane structure, therefore, must take into account these latter characteristics of the x-ray diagram and their variation from fiber to fiber.

A. *Orientation*

In most cellulose fibers such as cotton, ramie, and flax, the x-ray diffraction rings are concentrated into arcs to give a fiber pattern. This characteristic is directly related to the orientation or alignment of the cellulose crystallites (*b*-axes of unit cell) parallel to the fiber axis (48). There is also present in some membranes, such as *Valonia ventricosa* (2, 35, 49, 50), a further selective orienting tendency of the *101* plane parallel to the membrane surface, which causes the 101 line (innermost line in cellulose diagram) to be absent when the x-ray beam is perpendicular to the surface (38, 40).

B. *Width of diffraction ring*

For a given sample and slit system the x-ray diffraction lines should have a definite width. In most cellulose patterns this line breadth is greater than that theoretically expected. Furthermore, in many of the fibers which show orientation, the diffraction lines arising from planes parallel to the *b*-axis (equatorial lines) have a greater width than the lines arising from planes perpendicular to the *b*-axis (meridian lines) (19). The width of the equatorial lines also varies from fiber to fiber. In some of the bast or lignified fibers, such as ramie or wood, the lines are very broad; in cotton they have a medium width, while in *Valonia* they are very sharp (49).

C. *Amorphous pattern*

The presence of a certain amount of amorphous material in most fibers is indicated by a general fogging in the x-ray diagram. The nature of this fogging varies from fiber to fiber (47). In some fibers it approaches a broad amorphous band similar to that of a homogeneous liquid. In other fibers, such as wood, it takes the form of a disk extending to the central beam, which is typical of a heterogeneous amorphous material. The amount of fogging also varies from fiber to fiber. For example, in

cotton it is very faint, while in some of the woody or lignified tissues it may almost completely mask the crystalline cellulose pattern.

III. EFFECT OF VARIOUS TREATMENTS ON THE X-RAY DIAGRAM

Aside from the characteristics of the diffraction patterns themselves, other important information regarding cellulose membrane structure may also be obtained by observing how these characteristics are affected by various physical and chemical treatments of the membrane.

A. Purification

When fibers such as young cotton or wood, which give amorphous scattering, are subjected to the usual purification treatments, a large portion of the scattering is often removed without affecting the crystallinity or orientation of the cellulose (41, 42). In some cases there is also a sharpening of the diffraction lines (47).

B. Degradation

When cotton fibers, which give an oriented cellulose pattern, are treated with hydrochloric acid (15), they lose most of their original physical properties and may be ground to a powder. This powder gives the same crystalline x-ray diagram with slightly sharper lines, but the orientation is now random.

C. Swelling

On the basis of the change in the x-ray diagram, Katz (23) has classified swelling as either intermicellar or intramicellar. In the first case (intermicellar), such as produced by water, or weak acids and alkalies, there is no change in the cellulose pattern; an amorphous liquid pattern becomes superimposed upon the pattern of the original cellulose. In the second case (intramicellar), such as produced by strong alkalies or acids, there is a change in the crystalline cellulose pattern. This change may consist of the formation of an entirely new crystalline pattern, corresponding to the formation of a swelling compound between the cellulose and the reagent (e.g., strong sodium hydroxide), or it may consist of a complete disappearance of the cellulose pattern (e.g., strong sulfuric acid). In most cases where the swelling agent produces a new crystalline diagram, this diagram reverts to that of hydrate (mercerized) cellulose upon removal of the swelling agent.

D. Deformation

If a stress is applied to a swollen fiber a certain amount of viscous flow takes place and the fiber is extended. The only effect produced upon the x-ray diagram by this process of extension is to change the intensity dis-

tribution around the diffraction ring, corresponding to a change in the orientation of the crystalline cellulose material (28, 47).

E. Dispersion

Many of the swelling processes, such as cuprammonium and viscose, which produce a new x-ray diagram are also able to swell further and disperse the fiber. If the dispersed cellulose is later coagulated in the form of fibers, a definite crystalline x-ray pattern (hydrate) is obtained. Furthermore, when the fiber is extended there is again the unmistakable evidence of cellulose crystallites which move as a unit (28).

It is possible to cite many other processes which indicate a discontinuity of cellulose crystalline structure, but the above examples will suffice to illustrate the general nature of the x-ray data which need to be explained.

IV. THEORIES OF CRYSTALLITE STRUCTURE

In discussing the structure of a discontinuous cellulose membrane built up of crystalline and amorphous areas which will satisfy the previously outlined x-ray data, mention will be made of only three concepts. These will be referred to as the micellar theory, the continuous structure theory, and the cellulose particle theory. It is not the purpose of the present paper to review the literature or discuss in detail the relative merits of these concepts, but rather to outline briefly how they serve to explain the x-ray data. Special emphasis will be placed on the presentation of unpublished observations correlating the x-ray data with the microscopic behavior of the cellulose particle (12, 47).

A. Micellar theory

The micellar theory was first postulated by Nägeli (34) in 1858. His concept of submicroscopic crystalline particles was based upon polarized light and swelling studies. Later on, after 1920, when x-ray workers found cellulose to be definitely crystalline, the concept was again revived to describe the cellulose crystallite (33). According to Meyer and Mark (30, 32), these micelles are submicroscopic in size, are much longer than they are thick, and are arranged with their long axes roughly parallel or spiral with respect to the fiber axis. Primary valence forces hold the glucose units in the form of a chain; secondary valence forces hold the chains along side each other to form the micelle; and tertiary forces, or amorphous cementing material between the micelles, hold them together.

The estimated size of the micelle (over 500 Å. long and 50 Å. thick (19); see later discussion) would account for the abnormal line breadth of cellulose diffraction lines, while the arrangement of the micelles would account for the orientation effects obtained in the various cellulose fibers. The

presence of an intermicellar material, which was not present in Nägeli's original concept, would account for the amorphous portion of the x-ray diagrams. A hydrolyzing or oxidizing agent would tend to attack first the forces or material between the micelles, and thus account for the fact that these reagents may disintegrate the fiber without disintegrating the crystallite. When a fiber is treated with a swelling agent, the reagent may enter either between the micelles without affecting the crystalline cellulose (intermicellar swelling), or it may enter into the micelle, penetrate the spaces between the cellulose chains, and produce a new diffraction pattern (intramicellar swelling) (22, 23). If the fiber is stretched while in the swollen condition the micelles rotate or move as a unit to produce the improved orientation indicated in the x-ray diagram (28). If the fiber is treated with a dispersing agent, the solvating process begins with a swelling and subsequent dissolving of the intermicellar material, and if the conditions are right the micelles are dispersed without essentially altering their size.

B. Continuous structure theory

Since 1930 considerable emphasis has been placed on "long-chain" or "giant" molecules. Ultracentrifugal determinations (24, 52, 56), viscosity measurements (53, 54, 55), and related physicochemical methods (25) indicate the molecular cellulose chain to be more than ten times longer than the estimated size of the micelle. As a result there has evolved a new concept of the micelle, which will be referred to as the "continuous structure" theory (17, 26, 27, 36, 51, 54).

Although the x-ray data imply a discontinuous crystalline structure for cellulose, there is, however, no proof that the discontinuity consists of discrete crystalline particles or micelles as originally postulated by the micellar theory. The chains need not be broken lengthwise or separated sidewise to form individual crystals in order to explain the x-ray data. If it is assumed that the crystalline regularity of the cellulose chains is interrupted by regions where the chains are not sufficiently close or regular to form a crystal lattice, then these regions would produce in the x-ray diagram the same effect as though they were amorphous material separating well-defined crystallites. In the cell wall of a fiber those regions which possess a definite crystalline regularity of cellulose chains may be considered as "crystallites"; at other places, where the chains have an irregular arrangement, these areas may be considered as "amorphous" or "intercrystalline" cellulose material.

The various characteristics of the x-ray diagram may be explained upon this assumption in very much the same manner as with the original micellar theory. The irregular areas would account for the amorphous

portion of the diagram, while the size and arrangement of the crystalline areas would account for the line breadth and orientation of the crystalline pattern. When the fiber is swollen or dispersed the liquid would enter the meshwork of regular and irregular regions and separate the material into minute particles of cellulose. These particles, as they behave toward the x-rays and toward swelling and deformation, or as they exist in the dispersed or in the regenerated state, may be considered as micelles.

C. Cellulose particle theory

The original work of Farr and Eckerson (13, 14) and subsequent work by Farr (8, 9, 10, 11), showing the formation of cellulose membranes from microscopic particles of uniform size in linear arrangement, has led to the suggestion that the cellulose particle may be substituted for the micelle in the current interpretation of x-ray diffraction data (15, 39). These particles, which are ellipsoidal in shape (1.1 μ thick and 1.5 μ long), exhibit all the properties of crystalline cellulose and are covered with a coat of non-crystalline substance. The assumption that these particles are fundamental building units is based upon the fact that they may be observed in the living cytoplasm of young cotton fibers (13). During fiber growth the particles unite end to end to form fibrils, which are deposited in a spiral arrangement to build up the mature cell wall. Confirmation of their continued existence in the mature fiber comes from microscopic studies which show that the membrane may be disintegrated into fibrils and these, in turn, into particles by means of suitable solvents (14).

With these observations in mind, we turn now to a consideration of how this concept of membrane structure may serve as a basis for explaining the various characteristics of the x-ray diagram.

V. EXPLANATION OF CHARACTERISTICS OF CELLULOSE DIFFRACTION PATTERNS ON THE BASIS OF THE CELLULOSE PARTICLE

A. Orientation

In orientation studies it is desirable to work with single plant cells, just as single crystals are desirable in crystal structure determinations. One of the cells most satisfactory for orientation studies is the marine alga *Valonia ventricosa*, which often grows to a diameter of over 3 cm. This large size enables one to cut out small sections of the cell wall which may be mounted for x-ray examination (35, 38).

Microscopic studies (11) show that cellulose particles unite end to end at the inner surface of the growing cell to form fibrils. The mature membrane is built up of two sets of fibrils which lie in opposite directions, crossing each other at angles of approximately 80°. X-ray diagrams of the same

sample with the x-ray beam perpendicular to the membrane show the usual equatorial diffraction lines concentrated into four arcs, with two sets at approximately 80° to each other. Thus, by comparison of x-ray diagrams and photomicrographs it may be shown that the *b*-axes of the cellulose unit cells are oriented parallel to the axes of the fibrils, and hence to the long axis of the cellulose particle.

Further confirmation of this conclusion comes from another similar single-cell marine alga, *Halicystis.* The x-ray diagrams taken perpendicular to this membrane show random orientation (45). Microscopic studies show that the cellulose particles in *Halicystis* do not form fibrils as they do in *Valonia,* but exist separately in random arrangement (11). *Halicystis* also differs from *Valonia* in that the cellulose exists in the mercerized or hydrate form rather than in the usual native form (45).

The x-ray diagrams show a selective orientation for both *Valonia* and *Halicystis.* With the x-ray beam perpendicular, the 101 crystallographic line is missing, but with the beam parallel the 101 line now exists as two arcs. This means that the 101 crystallographic plane is oriented parallel to the membrane surface in addition to the orientation shown in the perpendicular diagram. Thus, *Valonia* may be described as having a double biaxial orientation, and *Halicystis* as a selective uniplanar orientation (40). This higher orientation cannot be detected in the microscope because of the ellipsoidal shape of the particle, but the microscope does not belie its existence, since cross sections show a laminated structure for both *Valonia* and *Halicystis* (11).

From comparative x-ray and microscopic studies on a membrane such as *Valonia,* which possesses a definite biaxial orientation, it is possible to draw certain conclusions regarding the structure of the cellulose particle (47). First, the cellulose chains (*b*-axis of the unit cell) run parallel to the long axis of the particle. Whether a single chain runs the entire length (1.5 μ) of the particle, or only part of the way, is not known. Second, the various crystallographic planes of the unit cell extend continuously in the same direction throughout the particle. In other words, the glucose units which build up the cellulose particle are arranged within the particle to form a crystalline structure which approaches that of a single crystal. From the intrinsic nature of the x-ray data it is impossible to say definitely that the cellulose particle does not contain still smaller crystalline units, but if such is the case the present data impose the condition that these smaller units have a perfect three-dimensional orientation within the particle.

By assuming the cellulose particle to be the cellulose crystallite, it has been possible in every membrane so far examined in our laboratory to account completely for the orientation observed in the x-ray diagram

on the basis of the cellulose particle orientation observed in the microscope (12, 47). For example, the high degree of fibering in the x-ray diagram of ramie fibers (38), and the rather wide deviation from a parallel orientation in cotton (42) may be directly correlated with the fibril orientation observed in the cell wall (43, 47). An interesting case is that of the *Avena* coleoptile, where bands of cellulose lie perpendicular to the long axis of the epidermal cells (16). X-ray diagrams of this material indicate that the cellulose should have a preferred orientation parallel to the long axis of the cell. Detailed microscopic studies show that these bands are made up of cellulose particles, united not end to end as they are in fibrils, but side by side with their long axes oriented in the direction indicated by the x-ray diagram.

B. Width of diffraction ring

The abnormal equatorial line breadth of cellulose diffraction patterns may be explained in several ways; it may be due to crystal size, strain, impurities, imperfectly formed crystals, or to a combination of any of these. From calculations based upon a theoretical relationship between crystal size and line breadth, Hengstenberg and Mark (19) have estimated the size of the micelle in ramie to be approximately 50 Å. in diameter and over 500 Å. long. The method has been employed with success in estimating the particle size of many colloidal materials, but the method is only valid provided other factors are absent (19). With cellulose fibers it is difficult to be sure that other factors are absent. For example, it is well known that a condition of strain within a crystal will produce a broadening of the diffraction lines (3), and the effect of a small amount of foreign atoms included in the crystal lattice in producing a broadening of the lines has been clearly demonstrated in the case of metals (6).

Since many non-cellulosic materials are so closely associated with cellulose during its formation in the cytoplasm, it is possible that there may be a sort of mixed crystallization of the cellulose with non-cellulosic materials at the surface of the particle. This is suggested by comparative x-ray and microscopic studies (12, 47) of a number of cellulose membranes which indicate that whenever a large amount of non-doubly refractive material may be observed on the surface of the cellulose particle, the equatorial diffraction lines are always broad and unresolved. If the surface of the cellulose particle, on the other hand, is observed to be comparatively free from non-doubly refractive material, then the diffracting lines are sharp and more clearly resolved. One would not expect the surface of the cellulose particle, molecularly speaking, to be as smooth as the surface of crystals which are built up of small molecules held together by secondary valences. It is not unreasonable to assume that the long primary-valence

cellulose chains would protrude at the surface and especially at the end of the particle to give a fringe structure. Such a structure would not only greatly increase the effective particle surface, but also permit a more permanent end-to-end union of the particles to form fibrils, and allow the polar hydroxyl groups of the protruding chains to become closely associated with other hydrophilic cytoplasmic constituents during particle formation. The outer portion of the particle would thus vary from crystalline cellulose to an amorphous mixture of cellulose and non-cellulosic materials at the outer surface. This outer portion would try to crystallize as best it could, producing a lattice strain and distortion which would extend throughout the particle.

If the broadening of the diffraction lines is due to impurities or strain, the question may arise as to why the equator and meridian lines are not equally affected. It will be remembered in this connection that the substitution of hydroxyl by other groups often affects only the crystallographic planes parallel to the cellulose chains, the meridian lines remaining constant (31).

C. Amorphous pattern

Comparative x-ray and microscopic studies (12, 47) indicate that the crystalline pattern of cellulose arises from the cellulose particle, which shows double refraction in polarized light. Furthermore, the size of the particle (1.1 x 1.5 μ) is small enough to give the Debye–Scherrer patterns observed. It is unnecessary to assume crystalline units smaller than the particle, since powder patterns are obtained whenever the crystals of the diffraction sample become smaller than about 5 μ.

There is also reason to believe that the amorphous pattern present in the x-ray diagram of cellulose fibers may be partly accounted for by the non-doubly refractive material which may be observed on the surface of the particles. This assumption is based largely upon purification studies. For example, the x-ray diffraction diagram of a young cotton fiber is very complex, showing the presence of both amorphous and crystalline non-cellulosic materials (41). This x-ray result is confirmed by microscopic examination (13). Extraction of these young fibers with organic solvents removes a fraction which gives an x-ray pattern characteristic of a waxy material; extraction with ammonium oxalate, a fraction which gives a pattern characteristic of a pectic material; and extraction with hot dilute alkali, a fraction which gives an amorphous pattern. After these extractions the residue gives only the crystalline diffraction pattern of cellulose, upon which there is superimposed a weak amorphous pattern. Microscopic examination shows this residue to contain cellulose particles covered with a thin coat of non-doubly refractive material, the other non-cellulosic materials originally present in the young fiber having been removed by

the purification process (41). In order to reduce further the intensity of the amorphous pattern it is necessary to give the residue such drastic treatment as to destroy the property of the cellulose particles to cement themselves together as a coherent membrane upon drying. These preliminary studies, therefore, would indicate that the amorphous portion of a cellulose membrane diffraction diagram may arise from two more or less overlapping sources. One source is that of the non-cellulosic materials which often may be removed by the usual purification processes without destroying the membrane structure. The other is that of an as yet unidentified material which is more intimately associated with the crystalline cellulose, and which apparently plays the rôle of a cementing material which holds the particles together.

VI. EXPLANATION OF THE EFFECT OF VARIOUS TREATMENTS ON THE BASIS OF THE CELLULOSE PARTICLE

A. Purification

Comparative x-ray and microscopic studies show that the amorphous pattern gets weaker and the cellulose lines sharper as the non-doubly refractive material, which may be observed on the surface of the particle, is removed by purification treatments (12, 47). It is difficult to picture this increased sharpness of the diffraction lines as being due to an increase in micellar size, a conclusion which must follow if the line breadth is attributed to the size of the micelle. It would seem more logical to conclude that the amorphous material exists as a separate phase, which is closely associated with the crystallinity of the cellulose.

B. Degradation

Since the cotton fiber may be disintegrated without destroying the crystalline x-ray diagram (15), this would indicate that the major portion of the cellulose exists in the form of crystallites which are less subject to attack than the intercrystalline material or forces which bind them together. Microscopic studies show that hydrochloric acid preferentially attacks the interparticle phase, with a resulting disintegration of the membrane into fibrils and these in turn into particles (15). The lack of change in the crystalline pattern does not necessarily mean that the chemical behavior of the cellulose chains has not been affected, but it does mean that the crystalline arrangement of the chains within the particle has not been disturbed.

C. Swelling

Comparative x-ray and microscopic studies on the same material (12, 47) indicate that in the case of intermicellar swelling, such as is produced by water or weak acids or alkalies, all of the swelling takes place in the

amorphous phase between the particles. The particle itself is not changed. This swelling or taking up of the liquid by the interparticle phase apparently accounts for the appearance in the x-ray diagram of the amorphous pattern which is superimposed upon the unchanged crystalline cellulose pattern.

With intramicellar swelling, such as produced by ethylenediamine (57), where there is the appearance of a new crystalline pattern corresponding to the formation of a swelling compound, the cellulose particle may be observed to increase in diameter. This increase in particle diameter is of the same order as the lateral increase in the unit cell dimensions calculated from the x-ray diagram (12, 47). The formation of a swelling compound does not destroy the double refraction of the cellulose particle, and after swelling has once taken place there is no further swelling of the particle or change in x-ray diagram with further application of the swelling agent. Intramicellar swelling is also produced by strong alkalies, but here it is accompanied by excessive fiber swelling and the appearance of a pronounced liquid pattern superimposed upon the crystalline pattern of the new swelling compound. In this case the liquid pattern and most of the fiber swelling may be accounted for by simultaneous swelling of the interparticle material (46).

In the second type of intramicellar swelling, such as produced by strong sulfuric acid, where the crystalline pattern disappears completely, microscopic examination shows the particle to swell first in diameter and later in all directions with the disappearance of double refraction (10). In this type the swelling is progressive with further addition of reagent until the particle is disrupted.

From the preliminary x-ray and microscopic results now available it would appear that the terms "inter- and intra-particle" swelling may be used synonymously with the terms "inter- and intra-micellar" swelling.

D. Deformation

Since the crystalline cellulose pattern changes only in orientation when a swollen fiber is stretched (28, 47), this would indicate that the crystallites themselves are not deformed, but that they move as a unit during the fiber deformation process. This is in keeping with comparative microscopic studies which indicate that the interparticle material, softened as a result of the swelling reagent, permits the fibrils and particles to glide past each other and move as a unit during deformation. All changes in orientation observed in the x-ray diagram may be accounted for by changes in orientation of cellulose particles (12, 47).

E. Dispersion

Since "regenerated" cellulose gives the same type of x-ray diagram (mercerized or hydrate) as cellulose which has been subjected to intra-

micellar swelling and shows the same evidence of crystallites which move as a unit when the fiber is extended, this would indicate that the crystallites resulting from these two processes are the same. This is in agreement with slit ultramicroscopic (5, 44) and dark-field microscopic studies (10), which indicate the presence of cellulose particles in Brownian movement in both cuprammonium and xanthate dispersions of cotton fibers. These reagents apparently first produce a swelling of the fiber (both inter- and intra-particle swelling) followed by a dispersion of the fiber to give cellulose particles in colloidal suspension (10, 44). In the presence of a coagulating bath the particles are flocculated and deswollen to give a product in which cellulose particles may still be observed. This mechanism of "solution" and "regeneration" would account for the fact that the x-ray diagrams of "regenerated" and "mercerized" cellulose are the same.

VII. DISCUSSION AND SUMMARY

The essential attribute of the micellar theory is the existence of discrete rod-like submicroscopic crystalline particles, which are oriented with respect to the fiber axis and are separated one from another by intermicellar material which coheres and yet allows the micelles to move as a unit during intermicellar swelling and plastic flow and also to be dispersed as a unit during the early stages of the solution process. The essential attribute of the continuous structure theory is the existence of long cellulose chains, which have crystallized out as best they can in the cell membrane in such a way that the crystalline regularity in the membrane is intercepted or warped by irregular regions which behave essentially as amorphous material toward x-rays and toward the penetration of swelling and dispersing liquids. The essential attribute of the cellulose particle theory is the existence of ellipsoidal microscopic crystalline cellulose particles, which are oriented with respect to the fiber axis and separated one from another by amorphous materials which cement the particles together and yet allow them to behave as a unit during swelling and plastic flow and also to be dispersed as a unit by many reagents.

With the exception of the size and shape of the particles, the essential features of the micellar and particle theory are identical. Furthermore, the interpretation of x-ray diffraction data on the basis of the two concepts is essentially the same. The interpretation differs, however, in that the identification of the cellulose crystallite with the micelle is founded upon the existence of a hypothetical building unit, while the cellulose particle is a unit whose existence may be demonstrated by direct experimental evidence,—microscopic observation. In identifying the cellulose crystallite with the particle, the only assumption necessary is one which will account for the width of the x-ray diffraction lines. If it is assumed that the particle has a continuous but imperfect crystalline

structure so as to account for the width of diffraction lines and the porosity of the particle to chemical reagents, then the major x-ray arguments for the existence of submicroscopic micelles in natural fibers cease to exist. The present preliminary data would indicate that the existence of crystalline cellulose particles separated by an amorphous interparticle material is adequate for explaining that portion of the x-ray data pertaining to cellulose membrane structure. This interpretation enables one to make a more direct correlation between x-ray and microscopic data without contradicting other physical and chemical data which point to the existence of a dual structure.

All of the chemical data which point to the preëxistence of well-defined micelles in native fibers rest upon acts of disruption and are therefore open to suspicion (1). The methods of viscosity (53, 54, 55), the ultracentrifuge (24, 25, 52, 56), diffusion coefficients (21), osmotic pressure (4), total surface (20), and methylation (18) are useful for estimating the average size of the "micelle" or "molecule" as it exists in solution, but unless one is certain concerning what happens during the solution process, it is difficult to extrapolate the results to untreated native cellulose membranes. The conclusion that it is unnecessary to assume a crystalline unit smaller than the cellulose particle in order to explain the x-ray data, however, in no way discredits the important rôle which cellulose chain molecules or a further secondary structure of the particle may play in the interpretation of many chemical and physical data. It merely emphasizes the fact that both microscopic and submicroscopic structures may play an important and often interrelated part in the behavior of native cellulose membranes.

REFERENCES

(1) Astbury, W. T.: Trans. Faraday Soc. **29**, 193 (1933).
(2) Astbury, W. T., Marwick, T. C., and Bernal, J. D.: Proc. Roy. Soc. (London) **B109**, 443 (1932).
(3) Barrett, C. S.: Metals & Alloys **5**, 154, 170, 196 (1934).
(4) Büchner, E. H., and Samwel, P. J. P.: Trans. Faraday Soc. **29**, 32 (1933).
(5) Compton, J.: Contrib. Boyce Thompson Inst. **10**, 57 (1938).
(6) Davey, W. P.: *Crystal Structure and its Applications.* McGraw-Hill Book Company, Inc., New York (1934).
(7) Eckling, K., and Kratky, O.: Z. physik. Chem. **B10**, 368 (1930).
(8) Farr, W. K.: J. Applied Phys. **8**, 228 (1937).
(9) Farr, W. K.: J. Phys. Chem. **41**, 987 (1937).
(10) Farr, W. K.: Contrib. Boyce Thompson Inst. **10**, 71 (1938).
(11) Farr, W. K.: J. Phys. Chem. **42**, 1113 (1938).
(12) Farr, W. K.: Unpublished results.
(13) Farr, W. K., and Eckerson, S. H.: Contrib. Boyce Thompson Inst. **6**, 189 (1934).

(14) Farr, W. K., and Eckerson, S. H.: Contrib. Boyce Thompson Inst. **6**, 309 (1934).
(15) Farr, W. K., and Sisson, W. A.: Contrib. Boyce Thompson Inst. **6**, 315 (1934).
(16) Farr, W. K., and Sisson, W. A.: Contrib. Boyce Thompson Inst. **10**, 127 (1939).
(17) Frey-Wyssling, A.: Kolloid-Z. **85**, 148 (1938).
(18) Haworth, W. N., and Machemer, H.: J. Chem. Soc. **1932**, 2270, 2372.
(19) Hengstenberg, J., and Mark, H.: Z. Krist. **69**, 271 (1928).
(20) Herzog, R. O.: Ber. **58B**, 1254 (1925).
(21) Hess, K.: *Chemie der Zellulose.* Akademische Verlagsgesellschaft, Leipzig (1927).
(22) Katz, J. R.: In K. Hess' *Chemie der Zellulose.* Akademische Verlagsgesellschaft, Leipzig (1927).
(23) Katz, J. R.: Trans. Faraday Soc. **29**, 279 (1933).
(24) Kraemer, E. O.: Ind. Eng. Chem. **30**, 1200 (1938).
(25) Kraemer, E. O., and Lansing, W. D.: J. Phys. Chem. **39**, 153 (1935).
(26) Kratky, O., and Mark, H.: Z. physik. Chem. **B36**, 129 (1937).
(27) Kratky, O., and Mark, H.: Papier-Fabr. **36**, 345 (1938).
(28) Mark, H.: Trans. Faraday Soc. **29**, 6 (1933).
(29) Mark, H.: Chem. Rev. **26**, 169 (1940).
(30) Mark, H., and Meyer, K. H.: Z. physik. Chem. **B2**, 115 (1929).
(31) Mathieu, M.: Trans. Faraday Soc. **29**, 122 (1933).
(32) Meyer, K., and Mark, H.: Ber. **B61**, 593 (1928).
(33) Meyer, K. H., and Mark, H.: *Der Aufbau der hochpolymeren organischen Naturstoffe.* Akademische Verlagsgesellschaft, Leipzig (1930).
(34) Nägeli, C. v.: *Die Stärkekörner.* F. Schulthub, Zurich (1858).
(35) Preston, R. D., and Astbury, W. T.: Proc. Roy. Soc. (London) **B122**, 76 (1937).
(36) Sauter, E.: Z. physik. Chem. **B35**, 117 (1937).
(37) Sisson, W. A.: Textile Research **4**, 286 (1934).
(38) Sisson, W. A.: Ind. Eng. Chem. **27**, 51 (1935).
(39) Sisson, W. A.: Textile Research **5**, 119 (1935).
(40) Sisson, W. A.: J. Phys. Chem. **40**, 343 (1936).
(41) Sisson, W. A.: Contrib. Boyce Thompson Inst. **8**, 389 (1937).
(42) Sisson, W. A.: Contrib. Boyce Thompson Inst. **9**, 239 (1938).
(43) Sisson, W. A.: Contrib. Boyce Thompson Inst. **9**, 381 (1938).
(44) Sisson, W. A.: Contrib. Boyce Thompson Inst. **10**, 113 (1938).
(45) Sisson, W. A.: Science **87**, 350 (1938).
(46) Sisson, W. A., and Saner, W. R.: J. Phys. Chem. **43**, 687 (1939).
(47) Sisson, W. A.: Unpublished results.
(48) Sisson, W. A., and Clark, G. L.: Ind. Eng. Chem., Anal. Ed. **5**, 296 (1933).
(49) Sponsler, O. L.: Nature **125**, 633 (1930).
(50) Sponsler, O. L.: Protoplasma **12**, 241 (1931).
(51) Sponsler, O. L.: Quart. Rev. Biol. **8**, 1 (1933).
(52) Stamm, A. J.: J. Am. Chem. Soc. **52**, 3047 (1930).
(53) Staudinger, H.: *Die hochmolekularen organischen Verbindungen.* J. Springer, Berlin (1932).
(54) Staudinger, H.: Ber. **70**, 2514 (1937).
(55) Staudinger, H.: Papier-Fabr. **36**, 473 (1938).
(56) Svedberg, T.: Ind. Eng. Chem., Anal. Ed. **10**, 113 (1938).
(57) Trogus, C., and Hess, K.: Z. physik. Chem. **B14**, 387 (1931).

THE CONTRIBUTION OF X-RAY RESEARCH TO THE KNOWLEDGE OF RUBBER[1]

S. D. GEHMAN

The Goodyear Tire and Rubber Company, Akron, Ohio

Received October 3, 1939

CONTENTS

I. INTRODUCTION

Ever since the observation by Katz, in 1924, that stretched rubber gave x-ray interferences similar to those for a crystalline material, the subject of x-ray diffraction by rubber has been one of continuous interest for research of a fundamental character dealing with the molecular structure of high polymers. At the start, this property of rubber was regarded as more or less individual and distinctive, but in the course of time the tendency has been to show that the behavior of rubber in this respect fits into the general picture of the x-ray diffraction by both natural and synthetic polymers. X-ray diffraction by rubber has thus lost much of its early character as an isolated phenomenon, and its full significance in connection with the general property of high elasticity now can be shown to lie in the quantitative measures of the diffraction and the circumstances, such as the temperature range and elongations, for which various types of diffraction occur.

[1] Presented at the Symposium on X-ray Studies of Substances of High Molecular Weight, held at the Ninety-eighth Meeting of the American Chemical Society in Boston, Massachusetts, September, 1939, under the auspices of the Divisions of Physical and Inorganic Chemistry, Colloid Chemistry, Organic Chemistry, Paint and Varnish Chemistry, and Rubber Chemistry of the American Chemical Society.

II. CRYSTALLINE AND AMORPHOUS UNSTRETCHED RUBBER

Physical measurements of density, heat capacity, and other quantities associated with changes of form have been interpreted as showing that raw rubber can exist in several modifications, depending upon the temperature. The x-ray patterns sharply distinguish only two modifications: amorphous (60), in which case a pattern similar to that of a liquid is secured, or crystalline, for which a system of Debye–Scherrer rings occurs (55, 40) (see figures 1 and 2). In attempting to fix the transition temperature for these patterns, the situation in regard to the temperature modifications of rubber must be briefly considered.

The duration of the freezing, the pressure, and possibly the temperature during freezing are definitely factors in determining the transition temperature for rubber. If rubber is frozen for a period of several days at a temperature lying roughly between 6°C. and −10°C., it thaws in the range from 6°C. to 16°C. (3).

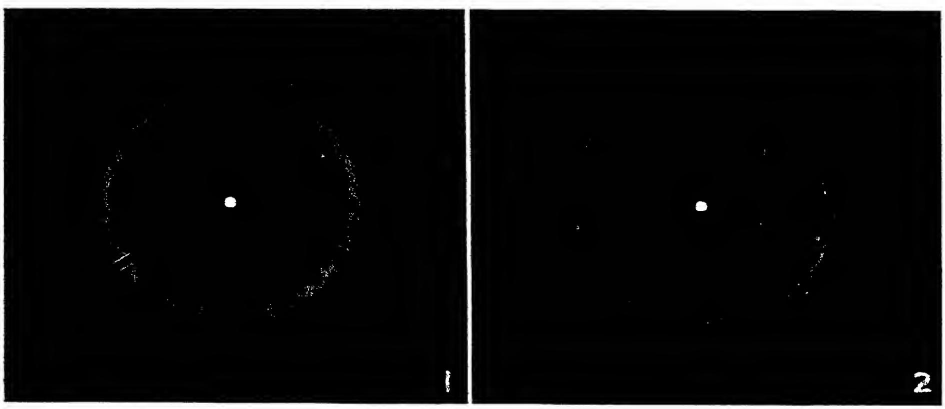


FIG. 1. Pattern for amorphous rubber
FIG. 2. Pattern for frozen rubber

For samples of rubber frozen under pressure, Thiessen and Kirsch (64) found that the crystalline phase was stable up to 20°C., placing the transition point, as determined by x-ray examination, between 20°C. and 23°C. When rubber is frozen for a period of years, it shows a still higher transition point. After being frozen for four years, a sample reported on by van Rossem and Lotichius (58) had a thawing point of 31–33°C.; nine years later this increased to 35–37°C. X-ray diffraction patterns were secured for some of these specimens by Katz (40). In the course of thawing, the Debye–Scherrer rings of frozen rubber gradually become more diffuse and lose intensity. The amorphous halo appears and increases in

intensity, its superposition causing a shift in the relative intensity of the Debye–Scherrer rings. Finally, the amorphous halo alone remains. These changes were studied photometrically by Katz. The transition region determined by the disappearance of the crystalline interferences coincided very closely with that determined by the physical measurements, the crystalline interferences becoming imperceptible shortly before the thawing was completed, as is to be expected.

Meyer, v. Susich, and Valkó (52) placed the x-ray transition point even higher, at 35–45°C. Whitby (68) reports a transition temperature, determined by physical means, of 43.5°C. for a sample of rubber believed to have been frozen for thirty years.

The fact that specimens of frozen rubber with thawing points ranging from 8°C to 43.5°C. all give the same x-ray interferences, differing perhaps in intensity and sharpness, indicates that these various forms represent stages in the exceedingly slow but continuous approach to a stable equilibrium condition rather than distinct modifications. Expressed in terms of the mechanism of crystallization in ordinary liquids, the molecular mobility in rubber at low temperatures must be very small, leading to a low velocity of crystallization. The other factor affecting the speed of crystallization is the rate of formation of nuclei. Owing to the interaction of these factors, the occurrence of crystallinity in rubber proceeds with reasonable speed in a temperature range limited to about +8°C. to −35°C. If the temperature is lowered through this range before the crystallization process is complete, which evidently requires a period of years, the rubber exists in a supercooled state, stable at the lower temperature for all practical purposes. Thus, Pummerer and v. Susich (57) found that rubber held at −190°C. for 8 days gave an amorphous x-ray diagram. This is a notable example of the lack of correlation which can exist between the x-ray diagram and the physical properties.

Bekkedahl (3) found a second transition for both crystalline and amorphous rubber at −72°C. This transition has not been reported in x-ray work, and it is doubtful if it would be perceptible except possibly by photometric measurements of the Debye–Scherrer rings or the halo, respectively.

High pressure has been reported to prevent the crystallization of rubber in the usually favorable range of temperature. In an experiment by Dow (16) crude rubber was held at a pressure of 8000 kg. per cm.^{2} at 0°C. for a period of 14 days, but no evidence of freezing was obtained. This was attributed to the high viscosity induced in the rubber by the pressure. Thiessen and Kirsch (64), on the other hand, working in a lower pressure range of 10 to 25 atmospheres, found that at a given temperature, the higher the pressure, the shorter the time required for crystallization. The application of pressure also enabled crystallization to occur at higher tem-

peratures than would have been otherwise possible. Under a pressure of 25 atmospheres crystallization took place at 10°C., but 15 atmospheres was sufficient to bring about crystallization only at 8°C., as determined by x-ray examination. Comparing the results of Dow with those of Thiessen and Wittstadt, there is evidently a favorable range of pressure for the crystallization of rubber as well as a favorable range of temperature.

The question now arises as to whether the crystalline or amorphous form of rubber is the stable one under ordinary conditions. The thawing points of old samples of frozen rubber are so high as to indicate that this is the stable form, although there is a possibility that the thawing points might be somewhat lower if the thawing process, as the freezing, were of long duration. Crystallization at room temperatures proceeds so slowly, presumably owing to the slow rate of formation of crystallization nuclei, that for all practical purposes the supercooled amorphous form is also stable. Thus amorphous patterns are the general rule for crude rubber under ordinary circumstances. A number of observers have reported the occurrence of crystalline interferences, especially for crepe rubber (42, 57), but it is doubtful if these represent crystallization at room temperature rather than a persistence of crystallization which occurred in the favorable range of temperature.

In general, unstretched vulcanized rubber gives the amorphous halo under all conditions. But vulcanization can be carried out to various degrees, and, presumably, lightly vulcanized samples might be induced to crystallize. The occurrence of the three-dimensional network of primary valence linkages in vulcanized rubber evidently interferes with the movement of the molecules into a crystal lattice to such an extent that crystallization cannot occur spontaneously. In contrast, rubber which has been milled or masticated freezes readily.

III. THE X-RAY DIFFRACTION PATTERN FOR AMORPHOUS RUBBER

In some of the earlier work, rubber was reported to give two halos corresponding to spacings of 6.03 Å. and 14.88 Å. (8, 28). Katz, however, consistently reported the pattern to be a single halo (36, 41). The above values were computed from a formula due to Ehrenfest, which gave somewhat higher values than the Bragg relation now generally used for this purpose. The larger spacing is not mentioned in later work, in which the pattern is regarded as consisting of a single halo corresponding to a spacing usually given as between 4.5 and 4.7 Å. (2, 10, 48). Mark and v. Susich, in photometering the pattern, found a weak halo corresponding to a spacing of 8 Å. This also has not been verified, and in view of the work of Simard and Warren (61), using strictly monochromatic radiation, it appears that these larger spacings which have been reported may have

been spurious effects due to continuous background or other causes. With filtered Cu K_α radiation a considerable amount of blackening usually occurs in the central region of the diagram. Further tests with monochromatic radiation are desirable to check this point.

Unstretched vulcanized rubber gives a halo which, as far as is known, is identical with that for unvulcanized rubber. There is some evidence, however, that the halo decreases in diameter as vulcanization proceeds to a stage approaching ebonite (43), showing the occurrence of slightly larger intermolecular spacings. Rubber which has become hard and brittle, owing to oxidation, is also amorphous. Purified sol and gel rubber both give the same halo characteristic of total rubber (12). In addition, for gel rubber Clark observed a large spacing of 54 Å.

The same pattern thus persists in the various forms of amorphous rubber and is indicative of a structure which must resemble that of a liquid in many respects. Katz (41) carried out extensive comparative studies of the patterns of liquids before and after polymerization, in order to reach a better understanding of the meaning of the pattern for rubber. In the case of isoprene and the butadienes, he found that the patterns of the monomers and polymers were identical. Thus, the long chain molecules formed during polymerization and existing in natural rubber assume a short-range structure similar to that in an ordinary liquid. This essentially liquid structure in unstretched, amorphous rubber is frequently lost sight of in attempts to explain the properties of rubber on a mechanical molecular basis. It is, however, one of the most certain features of the molecular structure in rubber.

Simard and Warren (61) have carried through for the pattern of unstretched rubber the same type of Fourier analysis which has been so useful for investigations of the structure of liquids and of glasses. This analysis gives directly from the x-ray scattering curve the radial distribution of neighboring atoms about any given atom without any assumptions in regard to the structure. The results showed that each carbon atom has on the average 1.98 carbon neighbors at a distance of 1.52 Å. Beyond these nearest neighbors are approximately 3.4 carbon atoms at an average distance of 2.68 Å. Further concentrations occur at about 4.0 and 5.0 Å. These results were considered to be in good agreement with a structure composed of long chain molecules, such as is now usually assumed for rubber, using accepted values of the bond lengths and angles. The first three concentrations represent successively remote neighboring atoms in the chain. The concentration at 5 Å. corresponds to an atom's nearest neighbors in other molecules and represents the fairly definite distance of closest approach for carbon atoms of different molecules, such as is observed for many organic liquids. This concentration gives rise to the main peak

in the diffraction pattern, and the above distance coincides with the spacing as determined by the application of Bragg's law.

It follows from this work that the diffraction pattern for amorphous rubber offers no features which can give information as to the relative orientation or form of the chain molecules (see figure 3). The superficial resemblance of the pattern to that for a liquid is thus borne out by the analysis. The liquid structure must be reconciled in some way with the existence of the long chain molecules, which are held responsible for the profound difference in the physical properties as compared to liquids. In

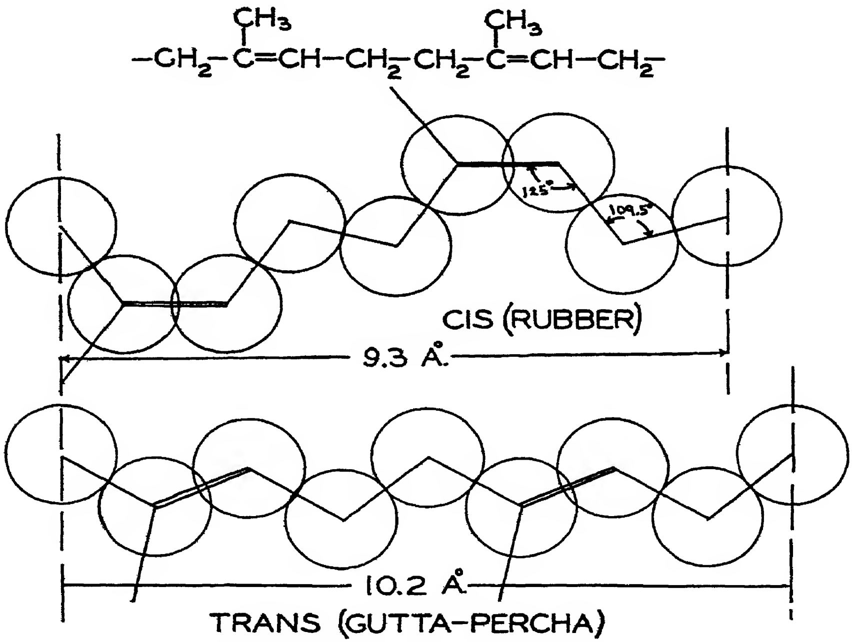


FIG. 3. Planar forms of the rubber chain molecule

the case of stretched vulcanized rubber, a tendency of the amorphous halo to split into two equatorial arcs has been observed (34, 43). This can be interpreted as evidence of the alignment of the long chain molecules in the direction of stretching.

IV. THE X-RAY DIFFRACTION PATTERN FOR FROZEN RUBBER

Independent measurements of the spacings for the diagram from frozen rubber are in good agreement, despite a wide variation in the type of specimens. Measurements by Barnes (2) on the patterns for two samples of rubber, one smoked and the other unsmoked sheet which had been frozen

over a period of years, and those by Clark, Wolthuis, and Smith (12) on patterns for purified sol, gel, and the total rubber are given in table 1. As no significant differences were found for sol, gel, and total rubber, the values reported for each have been averaged together. Older measurements by Ott (55) for frozen crepe rubber are also in fair agreement, although not as complete.

To all appearances the pattern for frozen rubber resembles that obtained from powdered organic crystals. It shows a structure of minute crystallites in random arrangement. Because of the insolubility of gel rubber,

TABLE 1

Spacings for frozen rubber

d*	INTENSITY*	d†	INTENSITY†
Å.		Å.	
6.26	s	6.18	s
		5.51	mw
4.97	vs	4.97	s
4.21	vs	4.24	vs
		4.13	
3.74	s	3.70	ms
3.41	m	3.46	m
3.02	wm	3.03	
2.78	w	2.82	
		2.73	w
2.58	m	2.55	
2.41	vw	2.38	vw
2.23	wm	2.21	
2.07	wm	2.07	
1.93	w	1.96	w
1.84	vw	1.83	w

* Measurements by Barnes (2).
† Measurements by Clark, Wolthuis, and Smith (12).

it might be suspected that this material contains primary valence cross-linkages. The crystallization observed may be interpreted as showing that any such network must be an extremely loose one.

V. THE X-RAY DIFFRACTION PATTERN FOR STRETCHED RUBBER

The most striking aspect of x-ray diffraction by rubber is the fact, first observed by Katz (36, 38, 39), that when rubber is stretched beyond a critical elongation it gives a fiber diagram with the spots arrayed as in the rotation diagram of a crystal, the axis of rotation corresponding to the direction of stretching (see figures 4 and 5). These diffraction spots appear superposed on the amorphous diagram. They increase in intensity, and

the halo decreases in intensity, as the stretching proceeds. The pattern has been substantiated for a wide variety of rubbers from different sources (25, 11). It also generally occurs for rubber purified by diffusion (56, 67), although the sol rubber prepared at the Bureau of Standards and investigated by Clark (12) did not show a fiber diagram at room temperature.

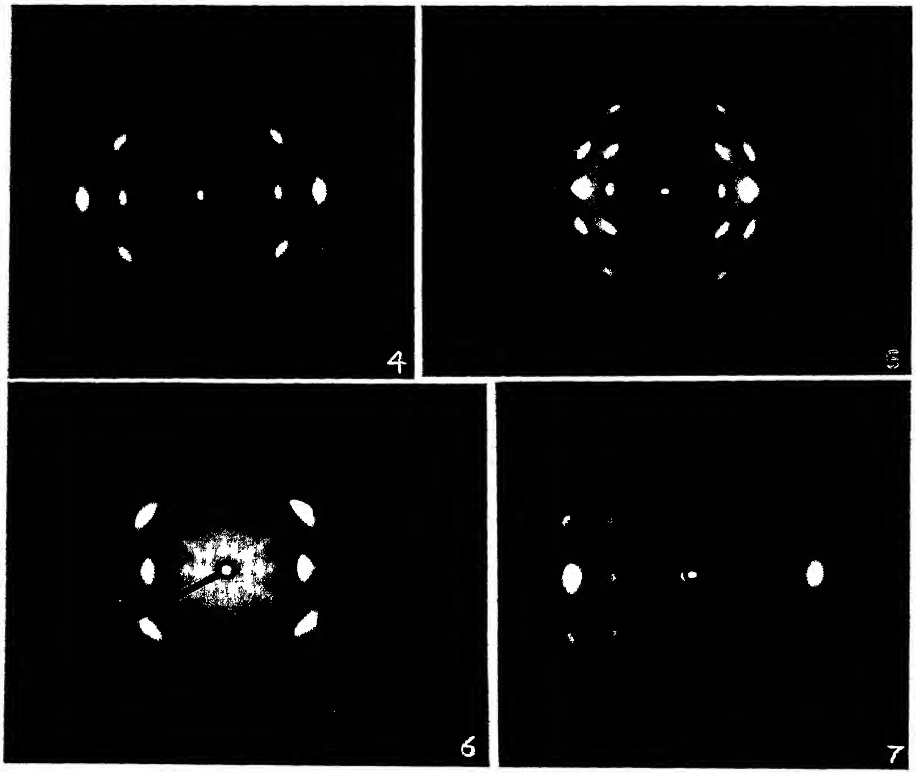


FIG. 4. Pattern for stretched rubber
FIG. 5. Pattern for racked rubber (cylindrical camera)
FIG. 6. Pattern for stretched rubber with higher orientation; beam perpendicular to plane of sheet
FIG. 7. Pattern for stretched rubber with higher orientation; beam parallel to width of sheet

For this material, presumably rendered too plastic by the purification process, a fiber diagram would probably have been obtained at a somewhat lower temperature. This view is supported by the fact that crystallization occurred both upon freezing and when the plasticity was reduced by vulcanization. Evidence is accumulating that the use of ethyl ether as the diffusion medium in the preparation of sol rubber is objectionable (22,

54, 44). Pummerer used ethyl ether but did not claim to observe any significant difference in the x-ray patterns of sol and gel rubber. Nevertheless, it is apparent from the text that some difficulty was experienced in securing the pattern for this sol rubber. When sol rubber is prepared by diffusion in petroleum ether, which may be regarded as a solvent representative of light hydrocarbons, fiber patterns are readily obtained at elongations comparable with those required for latex films. In view of these facts, any conclusions about the structure of rubber based on a qualitative difference between the x-ray behavior of the sol and gel fractions seem unwarranted.

The straightforward interpretation of such a fiber diagram is that crystallites are present which are aligned with one axis in the direction of stretching, the other two axes being in random orientation. It has been shown (55, 2, 12) that the lattice spacings for the Debye–Scherrer rings of frozen rubber agree exactly with those for stretched rubber, proving that the randomly oriented crystallites have the same crystal structure as those in stretched rubber. The lattice spacings depend upon the temperature to an extent determined by the thermal coefficients of expansion, which have been determined in this way (51).

The determination of the lattice in the crystallites and the structure of the unit cell from the fiber diagram has proved to be a difficult problem. Hauser and Mark (27) attempted to index the diffraction spots, but intensity relations discovered subsequently caused the early abandonment of these indices. Later attacks on the unit cell have followed the lines laid down in the investigation of Mark and v. Susich (48).

Mark and v. Susich showed that for stretched films of rubber characteristic changes in the intensity of the diffraction spots occurred, depending upon whether the direction of the x-ray beam was perpendicular or parallel to the plane of the film (see figures 6 and 7). Notably, the first equatorial spot was practically absent when the beam was parallel to the film. The second equatorial spot, on the other hand, was extinguished when the beam was perpendicular to the plane of the film. Rotation of the film gave the normal diagram. The conclusion drawn from these relationships was that the crystallites in the stretched films had all three axes aligned so as to give in effect the structure of a single crystal. Such specimens are said to show "higher orientation." The conditions of stretching necessary to produce higher orientation were studied by Gehman and Field (24). Higher orientation was found to occur when the stretching was carried out in such a way that the per cent contraction in gauge exceeded the per cent contraction in width. This can be readily accomplished by using short, wide sheets of rubber.

The interpretation of the results of Mark and v. Susich was that the

first two equatorial spots, designated as A_1 and A_2, arose from families of planes which were essentially perpendicular to each other. This led to the assumption of an orthorhombic lattice, the A_1 spot being indexed as (200) and the A_2 spot as (020). Four long-chain molecules were pictured as traversing the cell in the direction of the fiber axis. From the absence of diffraction by the planes (100), (010), and (001) in uneven orders it was concluded that twofold spiral or screw axes pass through the crystallites in all three directions, leading to the selection of the space group V^4. The twofold spiral axes were explained as arising from the *cis*-form of the hydrocarbon chain (see figure 8), alternate isoprene residues being turned at 180°.

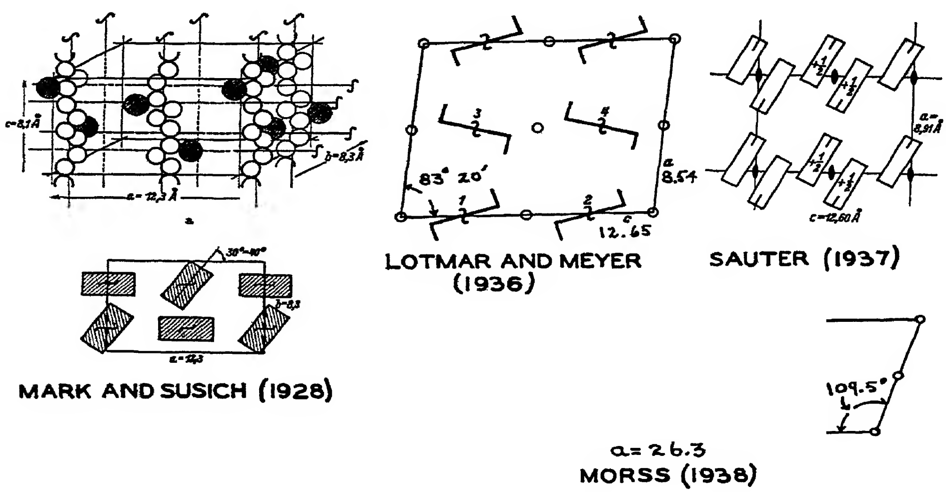


FIG. 8. Base planes of proposed unit cells

There were several discrepancies in this analysis. The density can be calculated from the unit cell by use of the formula

$$d = 1.65Mz$$

where d is the density in grams per cubic centimeter, M is the molecular weight of the repeating group in the chain molecule, z is the number of such groups in the unit cell, and V is the volume of the cell in Å.[3]

The density so calculated was about 1.08 as compared to the highest measured value, obtained by stretching and freezing, of 0.965. The A_4 equatorial reflection, assigned indices of (040), did not conform to the lattice. The fiber identity period was too small to accommodate the two isoprene groups of a zigzag planar hydrocarbon chain molecule such as was assumed.

Lotmar and Meyer (46), recognizing these difficulties, undertook a redetermination of the structure by means of a graphical evaluation of the fiber diagram. In this way the A_4 equatorial spot was indexed as (304), which served to determine the axial angle of a monoclinic cell. The twenty-three lattice spacings measured agreed well with those calculated for the cell. These results are given in table 2.

TABLE 2

X-ray interferences of stretched rubber (Lotmar and Meyer)

DESIGNATION OF SPOT	INDICES hkl	SPACING	INTENSITY	$\frac{\sin\theta}{\lambda}$ (MEASURED)	$\frac{\sin\theta}{\lambda}$ (CALCULATED)
		Å.			
A_1	(002)	6.29	s	0.079	0.079
A_2	(200)	4.24	vs	0.118	0.118
A_3	(004)	3.14	w	0.159	0.159
A_4	(304)	2.23	vw	0.224	0.224
I_1	(111)	5.50	vw	0.091	0.090
I_2	(012)	5.00	s	0.100	0.100
I_3	(210), (013), (211)	3.76	fs	0.133	0.133, 0.134, 0.134
I_4	(21$\bar{1}$), (212)	3.45	w	0.145	0.142, 0.147
I_5	(114)	2.86	vw	0.175	0.174
I_6	(312), (311), (214)	2.58	w	0.194	0 195, 0.195, 0.196
II_0	(020)	4.10	m	0.122	0.122
II_1	(120), (121)	3.70	w	0.135	0 135, 0.139
II_2	(12$\bar{1}$), (022)	3.45	m	0.145	0.143, 0.145
II_3	(220), (023), (221)	3.01	w	0.166	0.170, 0.170, 0.171
II_4	(22$\bar{1}$), (222)	2.78	w	0.180	0.177, 0.181
II_5	(22$\bar{2}$)	2.56	vw	0.195	0.193
II_6	(12$\bar{4}$)	2.39	vw	0.209	0.209
II_7	(224)	2.20	vw	0.227	0.223
II_8	(22$\bar{4}$)	2.08	w	0.240	0.242
II_9	(421), (420)	1.92	vw	0.260	0.264, 0.265
III_1	(130), (131)	2.58	vw	0.194	0.192, 0.195
III_2	(230), (231)	2.29	vw	0.218	0.217, 0.219
III_3	(034)	2.06	vw	0.243	0.242

The density calculated from the cell of Lotmar and Meyer is 1.02. This represents an improvement over the value from the cell of Mark and v. Susich, but it is still high. As shown by Smith and Saylor (62), a density of 1.02 for completely crystallized rubber implies that only 45 per cent of the rubber is crystallized in a sample of density 0.965, taking the density of amorphous rubber as 0.92. Although the presence of some amorphous rubber, even in highly stretched rubber, is indicated by a faint halo, such a proportion seems unreasonable unless the amorphous rubber

exists in a form which does not give rise to the characteristic halo. Meyer and Mark (50) judged that at least 80 per cent of the rubber in the extended form was crystallized.

Lotmar and Meyer assumed a spatial or three-dimensional chain form, the exact shape of which was considered to be not yet definitely established. For the cell proposed, four primary valence chains passed through the cell in the direction of the fiber axis, each a spiral with a definite direction of rotation due to the positioning of the isoprene residues. The chains were considered to be alternately stereoisomeric right-hand and left-hand chains. The reflections and extinctions led to the deduction of the space group C_{2h}^5.

Sauter (59) claimed to show that the unit cell of Lotmar and Meyer was impossible. Working with racked samples showing higher orientation, the A_4 interference did not occur at an angle of about 45° to the direction of the A_1 and A_3 reflections in his goniometer pattern obtained when the x-ray beam traversed the sample in the direction of the fiber axis. But the interferences for the axial pattern were long arcs, and Mark and Meyer (47) have disputed his contention. Sauter deduced an orthorhombic unit cell. He considered that the strong equatorial A_2 spot was a double spot, which he indexed as (002), (300). This results in a larger volume for the unit cell and a calculated density of 0.974. Sauter did not give indices except for the equatorial spots, so that the fit of the cell for all the interferences was not submitted.

Morss (53) has reported measurements of about forty lattice spacings, obtained with a cylindrical camera. The results are expressed in coördinates of the reciprocal lattice and greater relative than absolute accuracy was claimed. The ξ values are larger than those calculated from the data of Lotmar and Meyer or Sauter. The cell of Lotmar and Meyer fitted this data well, with the exception of two interferences for which the discrepancy appeared somewhat greater than the probable error of measurement. The planar indices used in the comparison of the cells are not given. Morss regarded this discrepancy as sufficient reason for the rejection of the cell of Lotmar and Meyer. He proposed a monoclinic cell of his own, which, using the data of Lotmar and Meyer to secure absolute values of the lattice spacings, gives a calculated density of 0.965. Morse considered that the work of Mark and v. Susich did not establish that A_1 and A_2 were diffractions by perpendicular families of planes.

These various unit cells are compared in table 3. The base planes of these cells, containing the axes perpendicular to the fiber axis, are repre sented in figure 8.

The unit cell of Mark and v. Susich has obviously been the starting point for the deduction of all these cells. The question remains as to

whether these attempts at refinement have yet reached the truth. It is very doubtful if the unit cell proposed by Sauter can be maintained. It is based on the proposition that the strong equatorial spot A_2 is a double spot. No photometric evidence for this has been offered. Furthermore, it is possible to prepare samples showing higher orientation to such a degree that the intensity of the A_2 spot, (002) by Sauter's indices, can be reduced to a value less than 1 per cent. of its normal value as compared to A_1. This small residual intensity, even if due entirely to a (300) interference, would not be sufficient to change the measurement of the spacing for the A_2 spot to the extent required for the Sauter cell.

The use of a monoclinic cell, suggested by Lotmar and Meyer, has undoubtedly brought about an improved agreement with the data. This is to be expected, since there are so many more interferences from which to choose for a monoclinic as compared to an orthorhombic cell, and the number of diffraction spots in the diagram is so limited. Critical tests such as those proposed by Sauter are therefore especially desirable for a monoclinic cell, and the outcome of a number of such independent experiments may very well be capable of giving a definite verdict. In this connection the diagram shown in figure 9 is of interest. This was obtained with the x-ray beam passing in the direction of stretch through a frozen sample showing higher orientation. A more elegant technique for securing suitable samples for this experiment is to vulcanize the stretched sample in sulfur chloride vapor. The structure is retained when the sample is removed from the clamps. The A_2 interference in figure 9 shows two distinct maxima at an angle of about 20° with the equator. The simplest interpretation of this is to regard it as evidence for a monoclinic cell. Accurate measurement of the angle is difficult, owing to the fact that the interferences still overlap in the best patterns which we have been able to obtain.

TABLE 3

Comparison of various unit cells proposed for stretched rubber

SPONSOR	CLASS	DIMENSIONS			MONOCLINIC ANGLE	CALCULATED DENSITY
		a	*b*	*c*		
		Å.	Å.	Å.		
Mark and v. Susich.	Orthorhombic	12.3	8 3	8.1 (f.p.)		1 08
Lotmar and Meyer.	Monoclinic	8.54	8.20 (f.p.)	12.65	83° 20′	1.02
Sauter . . .	Orthorhombic	8.91	8 20 (f.p.)	12.60		0.974
Morss . . .	Monoclinic	26.3	8.15 (f.p.)	8.9	109° 30′	0.965

The artificial aspects of fitting a cell to the data are well brought out by the investigation of Morss. A monoclinic cell arrived at by a systematic

variation of the monoclinic angle in a range of values from 90° to 115° was consistent with the diffraction data but was not considered really satisfactory because of the large number of possible diffractions as compared to the moderate number of spots observed. A smaller orthorhombic cell was found in the same investigation which accounted for all but one weak spot.

This apparent inadequacy of the diffraction data to lead to the unambiguous deduction of a unit cell may ultimately be resolved by independent

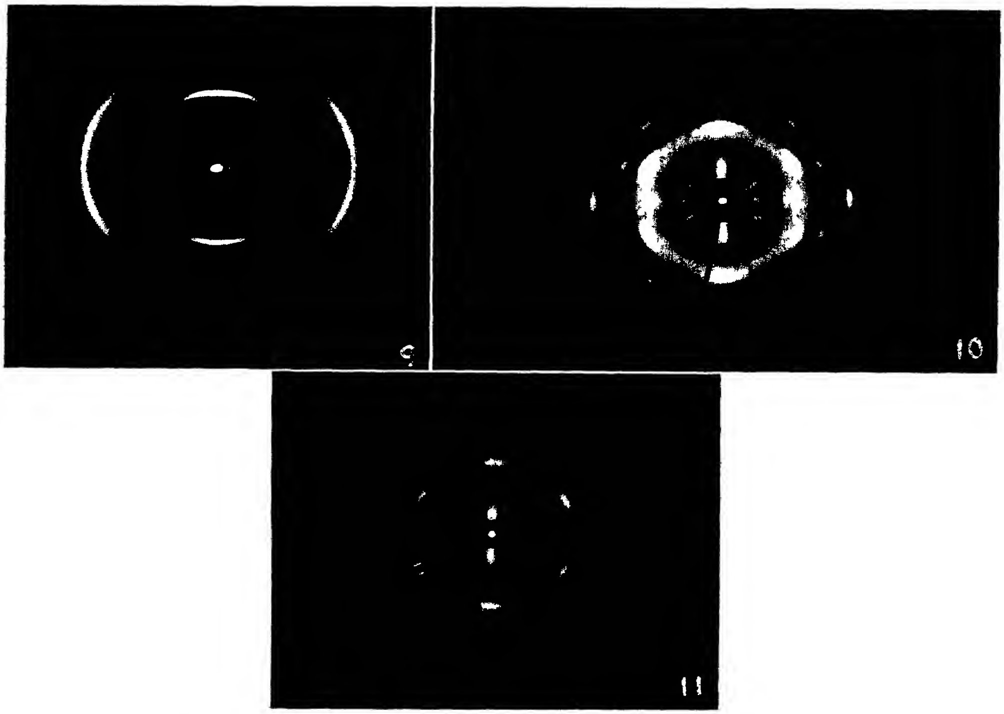


FIG. 9. Frozen sample; higher orientation; x-ray beam parallel to direction of stretch
FIG. 10. Pattern for racked rubber hydrochloride (cylindrical camera)
FIG. 11. Pattern for racked rubber hydrobromide

information, such as an improved knowledge of the structure of chain molecules. After all, the fitting of the diffraction spots to a space lattice is not an end in itself but is only one step in determining the structure of the chain molecules and their mutual packing. If more definite limitations in these respects could be imposed, the selection of the cell might become unique.

The value of the fiber period has already shown that the chain molecules in the crystallites of stretched rubber are not planar, a conclusion on

which there now seems to be general agreement. Sauter (59) has discussed various chain forms and the possibility that the symmetry of twofold spiral axes, such as is indicated by the extinctions of the diagram, exists not in the individual valence chains but because the chains are displaced by half the fiber period with respect to each other in such a way as to result in the densest packing. He has pointed out that the chain form is determined by the primary-valence field of force and the lattice forces. On this ground he regarded the digonal chain molecules as improbable. Comparative studies of chain forms will undoubtedly be of great assistance in determining the structure of the unit cell for rubber. Lotmar and Meyer (46) called attention to the fact that, owing to the possibility of rotation about the single bonds, the chain structure is not definitely established from the x-ray data even if recognized atomic distances and valence angles are retained.

TABLE 4

X-ray diffraction data for racked rubber hydrobromide

$\lambda = 1.54$ Å.

SPOT	θ	SPACING	INTENSITY
	degrees	*Å.*	
A_1	7° 3.5′	6.27	m
A_2	8° 8′	5.44	m
A_3	14° 15′	3.12	w
I_1	9° 26′	4.70	s
I_2	11° 58′	3.71	w
I_3	15° 54′	2.81	vw
II_0	9° 38′	4.60	vs
II_1	12° 48′	3.48	w

Another possibility for securing new information which may have a bearing on the chain form of rubber and the structure of the unit cell is a comparative study of the x-ray structure of various crystalline derivatives of rubber. Rubber hydrochloride, for instance, gives a crystalline diagram even without stretching, which passes over into an oriented fiber diagram when the material is stretched (23). Rubber hydrobromide shows the same type of structure. Patterns for these materials are reproduced in figures 10 and 11. The fiber identity period in both cases is 9.1 Å. An orthorhombic unit cell with *a* and *c* dimensions equal to twice the lattice spacings of the first two equatorial spots gives in each case fair agreement with the diffraction data and the density, assuming eight repeating groups to be present in the unit cell as for rubber. In the case of the rubber hydrohalides, preparations showing higher orientation have not been obtained. Thus the unit cell is more uncertain than it is for rubber.

Lattice spacings measured for rubber hydrobromide are given in table 4.

The measured density of the unstretched sample (bromine content, 47.8 per cent; theoretical, 53.6 per cent) was 1.51.

Although the addition of hydrohalides causes an appreciable lengthening of the fiber period as compared to rubber, the period is still too short to accommodate a planar zigzag chain with normal valence angles and carbon–carbon distances. The values of the fiber periods are compared in table 5. The angle between the double and single bonds was taken as 125° and that between single bonds as 109° 30′; the bond lengths were taken as 1.54 Å. for single bonds and 1.38 Å. for double bonds (5, 53).

In all these cases a characteristic shortening of the chains exists, so that the question of chain structure is involved in each case for the structure of the unit cell. Current theories of rubber elasticity postulate a statistical shortening of the chain molecules in the unstretched rubber and a straightening of the chains in the stretched condition as the fundamental basis of rubber-like elasticity. Thus, although the chain molecules in the crystallites are somewhat shortened when compared with normal chain forms, nevertheless they would probably have to be regarded as lengthened if compared to a most probable statistical length as determined by the possibilities of rotation about the single bonds.

TABLE 5

Comparison of calculated and measured fiber periods

	FIBER PERIODS CALCULATED FROM PLANAR CHAINS, NORMAL VALENCE ANGLES, AND BOND DISTANCES	FIBER PERIODS FROM X-RAY DIAGRAMS
	Å.	Å.
Rubber	9.3	8.2
Rubber hydrochloride	10.0	9.1
Rubber hydrobromide	10.0	9.1
β-Gutta-percha (20, 30)	10.2	9.54
μ-Polychloroprene (7, 21)	10.2 (*trans*)	9.6

VI. NATURE OF THE CRYSTALLITES IN STRETCHED RUBBER

Hauser and Mark (27) carried out an early investigation which disclosed a number of fundamental facts concerning the nature of the crystalline structure formed in rubber by stretching. As had already been noted by Katz (38), the fiber diagram appeared only if a minimum elongation were exceeded. For evaporated latex sheet, Hauser and Mark found this minimum elongation to be about 80 per cent, agreeing with the value given by Katz. There is every reason to suppose that this minimum elongation depends to some extent on the sample of rubber and on the room temperature. Much higher elongations, 150 to 300 per cent, are

usually necessary at temperatures of about 20–25°C. Above the minimum elongation, the intensity of the spots was found to increase approximately linearly with the extension (27, 49), and the amorphous halo to decrease linearly from the start of the stretching. Hock (33) showed that the heat evolved during extension was also proportional to the elongation above a minimum value.

These intensity relations and the conception of a sharp minimum elongation at which the diffraction spots occur for uncured rubber are due to an experimental error. This has been very unfortunte, since the linear relation of the intensity of the spots to the elongation has been difficult to understand and has led to a considerable amount of unsound theorizing.

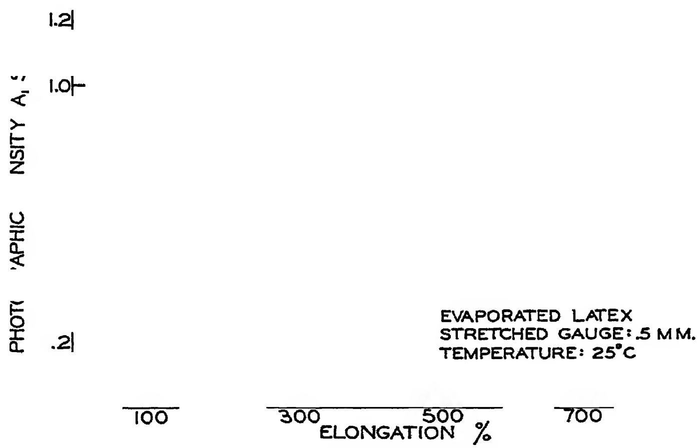


Fig. 12. Intensity of x-ray diffraction and elongation

In the experiments of Hauser and Mark no correction was applied for the diminishing thickness of the rubber as the sample was stretched. If samples of constant final thickness are used, the intensity of the spots is found to be related to the elongation by an S-curve such as is shown in figure 12. This implies that the crystallites begin to form from the start of the stretching and do not suddenly make an appearance. The form of the intensity–elongation curve is thus brought into closer correlation with the stress–strain curve. A study of the shape of the curve under various circumstances for both cured and uncured rubber would be an interesting investigation.

The positions of the interference maxima were found by Hauser and Mark to be independent of the extension within an accuracy of about 0.5 per cent. The "half-value width" of the spots did not change appreciably

during stretching. The stretching was therefore considered to be accompanied by an increase in the number of crystallites but not in their size. An improvement in the lattice perfection might also contribute to the greater intensity at higher elongations. Hengstenberg and Mark (31) made precise measurements of the width at half-intensity of the diffraction spots and, using the formula of Laue, arrived at dimensions of width, thickness, and length of 500, 150, and 600 Å., respectively, for the crystallites. No effect of temperature on the size of the crystallites could be measured (51). Diffuseness of the spots due to lattice imperfections would lead to small estimates of size. Most probably, the crystallites show a statistical distribution in size. Beset with uncertainties as calculations of size from the "half-value widths" are, these dimensions are, nevertheless, so much smaller than the molecular chain lengths now believed to exist in rubber, 7000 Å. or more, that the conclusion is fairly certain that a single long-chain molecule may pass through a considerable number of crystallites. For this reason, a secondary or micellar structure, such as was arrived at for natural cellulose fibers (19), has been presumed to exist in rubber (13, 14). It is untenable, however, that any such structure is due to botanical processes, since synthetic materials such as commercial polychloroprene (7, 21) and polyisobutylene (4) show essentially the same type of x-ray diffraction phenomena as natural rubber (see figures 13 and 14). Fisher and Gercke (18) have stated a number of other valid reasons for questioning the occurrence of a micellar structure in rubber. There is the possibility that the ordered regions in stretched rubber represent, in a sense, an unbroken continuity between the liquid structure and the structure of stretched rubber. Thus, comparatively short adjacent portions of the long chain molecules may assume relative orientations such as are now thought to occur for the individual molecules of an ordinary liquid. The formation of an actual space lattice would be favored by the relative immobility of the long chains, which can be thought of as traversing a number of the ordered regions.

As with frozen rubber, stretched raw rubber exhibits a series of roentgenographically determined melting points depending upon the extent to which crystallization has occurred. Thus, v. Susich (63) plotted a curve of the x-ray fusion point as a function of the elongation for raw rubber. The curve rose from a value of 20°C. at 150 per cent elongation to 90°C. at the highest elongations attainable in the experiments, about 700 per cent. Similarly, Feuchter and Hauser (17), in the course of a comprehensive x-ray study of racked rubber, showed that rubber racked to 600 per cent elongation had a "fusion point" or temperature at which retraction occurred at about 23°C., whereas, when it was racked to 1200 per cent, the fusion point occurred in the range of 35–45°C.

When raw rubber is stretched to intermediate elongations and frozen, crystallization occurs in such a way as to result, not in Debye–Scherrer rings, but in a strong fiber diagram (24). The crystallization proceeds from imperfectly oriented nuclei formed by the stretching. Clark (9) subsequently has published patterns taken under similar circumstances. As the secondary elongation during freezing is only about 5 per cent, this is the maximum elongation which can be thought of as occurring under these circumstances, due to a straightening and alignment of the long chain molecules from the most probable statistical form to the less probable ordered form in the stretched and frozen rubber. Unstretched rubber showing calender effect also gives an oriented diagram, especially when frozen (15, 43). Milling or mastication of crude rubber has the effect of progressively destroying the x-ray interferences upon stretching (27). Deterioration of vulcanized rubber by heating in air also inhibits the appearance of the fiber diagram (35). In the case of milling, it is interesting that the spots apparently diminish in intensity without noticeable increase in breadth, as would be the case if mastication had the effect of diminishing the size of the crystallites. This result falls into line with the views of the crystallite structure previously discussed. Exposure of a stretched piece of rubber to solvent also results in a diminution of intensity of the crystalline interferences (27). If masticated rubber is frozen and then stretched, a mixed Debye–Scherrer and spot pattern is obtained, as is the case for unmasticated, frozen, smoked sheet (28).

When rubber is vulcanized in a stretched condition at a temperature such that no crystalline structure exists, the resulting structure after vulcanization is also amorphous (29). In the case of cold vulcanization, a permanent fiber structure can be formed by vulcanization under stretch (6).

VII. TIME EFFECTS WITH THE X-RAY DIAGRAM OF STRETCHED RUBBER

Time is a factor in many of the x-ray diffraction phenomena with stretched rubber. The plastic properties are a reflection of internal molecular adjustments which require appreciable time intervals within which to be accomplished. Early in the x-ray work, Hauser, Hünemörder, and Rosbaud (26) found that smoked sheet could be stretched slowly to give an amorphous diagram. More recently, Hintenberger and Neumann (32) have published curves for uncured rubber relating the stress and elongation to the time. They found a well-defined maximum elongation and, in addition, a critical range of loading at comparatively low stresses, within which the rubber test pieces broke, when the elongation was plotted against the stress, the time of loading being constant. X-ray diffraction patterns offered an explanation for these interesting results. For the

relatively small stresses below the critical range, the elongation was attained so slowly that an amorphous pattern was secured; for samples stressed above the critical range, a strong fiber diagram resulted. Hence the structures in the stretched samples on either side of the critical range were entirely different. Crystallization is apparently necessary to withstand the higher stresses and to produce the ultimate strength generally associated with rubber.

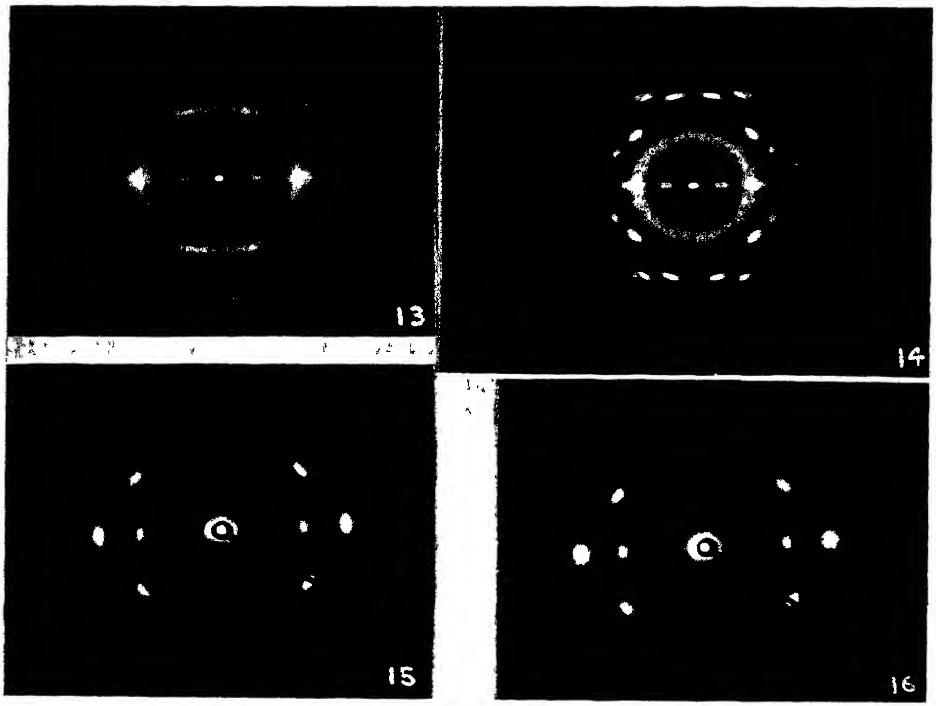


FIG. 13. Stretched polychloroprene (cylindrical camera)
FIG. 14. Pattern for stretched polyisobutylene (cylindrical camera)
FIG. 15. Pattern for a low-modulus gum stock
FIG. 16. Pattern for a high-modulus gum stock

Experiments showing x-ray hysteresis effects for vulcanized rubber were reported by Iguchi and Schoszberger (34). Among other hysteresis effects they found a large effect on the minimum elongation for fibering, depending upon whether the rubber was stretched or relaxed in a continuous or stepwise fashion. Thus, with slow relaxation, crystalline interferences were detected at as low an elongation as 130 per cent, whereas direct stretching required a minimum elongation of about 250 per cent for crystalline interferences.

Katz (37) and also Hauser and Mark (27) found that in some cases crystalline interferences persisted for appreciable periods of time with gradually diminishing intensity after the release of raw rubber which was stretched for a long time. This effect would undoubtedly be quite sensitive to the temperature and can be readily related to the slow approach to a stable equilibrium which rubber shows in practically all its crystallization processes.

A systematic study of the time lag of fibering for a vulcanized rubber compound was carried out by Davy and his associates (1, 45). The rubber underwent a standard stretching cycle and the x-ray beam, by suitable mechanical arrangements, passed through the test piece at definite time intervals after the stretching was completed. Time lags of the order of a second were observed for fibering to occur under the circumstances of the experiments. It would be interesting if such time lags could be related to the extremely variable plastic characteristics of vulcanizates, such as permanent set, creep or drift under load, relaxation of stress at constant elongation, and hysteresis.

Thiessen and Wittstadt (64, 65, 66, 69) observed that the x-ray fiber pattern continued to increase in sharpness and intensity for an appreciable time interval after stretching, an effect which was also investigated by means of double refraction. The change in anisotropy after termination of the stretching was termed "spontaneous orientation." Measurements showed that the orientation tended toward a definite equilibrium depending upon the temperature.

The molecular basis for the numerous hysteresis and drift phenomena with rubber is made evident by the various x-ray phenomena. The continuous slow adjustment of the molecules to new external circumstances of stress and temperature is an outstanding characteristic of the rubber structure.

VIII. THE X-RAY DIAGRAM AND THE PHYSICAL PROPERTIES OF VULCANIZED RUBBER

Since the same x-ray interferences on stretching are invariably obtained for rubber vulcanized under a wide variety of circumstances, it becomes necessary to examine the diagram in various secondary, quantitative ways to look for any information on the molecular structure built up by vulcanization. There are a number of possibilities for such measurements which have scarcely been touched in any serious effort to relate the x-ray diagram to the wide range of properties which can be secured by vulcanization. Katz and Bing (43) carried out an early investigation on the effect of vulcanization on the x-ray diagram, using much higher sulfur contents than is present compounding practice where accelerators are used. Katz

and Bing noted that progressive vulcanization tended to necessitate higher elongations for the appearance of the fiber diagram and to render the spots more diffuse. For present-day compounds, the intensity of the pattern also usually decreases somewhat for a technical range of cures as the cure progresses.

The statement is frequently repeated in the literature that the spots appear at an elongation of about 80 per cent for raw rubber and at 250 per cent for cured rubber. As previously stated, such a comparison is entirely a matter of temperature and of the extent of the vulcanization. At temperatures of 20–25°C. higher elongations are usually required for raw rubber. Furthermore, in most cases the comparison is not directly between uncured and cured rubber but between uncured rubber and rubber which has been masticated and subsequently cured. For cured and uncured latex compounds there is comparatively little difference in the intensity of the spots at low elongations such as 300 per cent. In the case of milled compounds, vulcanization does not act to inhibit the appearance of the pattern as is implied by stating that a higher elongation is required. Vulcanization, in this case, may make the pattern possible at room temperature. This state of affairs indicates that the network structure of the vulcanizate assists the crystallization by the same process of molecular immobilization which makes the rubber less sensitive to temperature. This assistance, however, extends only to the point where the rigidity introduced by the network begins to interfere with the movement of the molecules into the crystal lattice. If the network becomes too tight, crystallization is entirely prevented, as in the case of ebonite.

A study of the relation of the intensity of the diffraction spots to the elongation thus affords a possible means for investigating the structure of various vulcanizates. Other indirect experimental methods, such as the effect of temperature on the intensity at different elongations, also are available. The size, shape, and perfection of the crystallites and their orientation may also vary from stock to stock, as is suggested by the different appearance of the spots in the patterns shown in figures 15 and 16, for which the accelerators used were mercaptobenzothiazole and piperidinium pentamethylene dithiocarbamate, respectively. Thus, in spite of the fundamental limitations of the x-ray method and its failure to disclose the length of the chain molecules and dimensions of the network in the vulcanizate, nevertheless a number of indirect approaches remain open for the extension of the work to the problems of vulcanization.

Grateful acknowledgment is made of the assistance of Dr. J. E. Field in the experimental work.

REFERENCES

(1) Acken, M. F., Singer, W. E., and Davy, W. P.: Ind. Eng. Chem. **24**, 54 (1932).
(2) Barnes, W. H.: Can. J. Research **15**, 156 (1937).
(3) Bekkedahl, N.: J. Research Natl. Bur. Standards **13**, 411 (1934).
(4) Brill, R., and Halle, F.: Naturwissenschaften **26**, 12 (1938).
(5) Brockway, L. O.: Rev. Modern Phys. **8**, 231 (1936).
(6) Busse, W. F.: U. S. patent 1,909,455.
(7) Carothers, W. H., Williams, I., Collins, A. M., and Kirby, J. E.: J. Am. Chem. Soc. **53**, 4203 (1931).
(8) Clark, G. L.: Ind. Eng. Chem. **21**, 128 (1929).
(9) Clark, G. L.: Ind. Eng. Chem. **31**, 1397 (1939).
(10) Clark, G. L., Gross, S. T., and Smith, W. H.: J. Research Natl. Bur. Standards **22**, 105 (1939).
(11) Clark, G. L., Gross, S. T., and Smith, W. H.: J. Research Natl. Bur. Standards **23**, 1 (1939).
(12) Clark, G. L., Wolthuis, E., and Smith, W. H.: J. Research Natl. Bur. Standards **19**, 479 (1937).
(13) Clews, C. J. B., and Schossberger, F.: Proc. Roy. Soc. (London) **A164**, 491 (1938).
(14) Clews, C. J. B.: Proc. Rubber Tech. Conference, London, 1938, p. 955.
(15) de Visser, W.: *The Calender Effect and the Shrinking Effect of Unvulcanized Rubber*. Crosby Lockwood & Son, London (1926).
(16) Dow, R. B.: J. Chem. Phys. **7**, 201 (1939).
(17) Feuchter, H., and Hauser, E. A.: Kautschuk **5**, 194, 218, 245, 276 (1929).
(18) Fisher, H. L., and Gercke, R. H.: In Davis and Blake's *Chemistry and Technology of Rubber*. Reinhold Publishing Corporation, New York (1937).
(19) Frey-Wyssling, A.: Kolloid-Z. **85**, 148 (1938).
(20) Fuller, C. S.: Ind. Eng. Chem. **28**, 907 (1936).
(21) Garbsch, P., and v. Susich, G.: Kautschuk **8**, 122 (1932).
(22) Gehman, S. D., and Field, J. E.: Ind. Eng. Chem. **30**, 1931 (1938).
(23) Gehman, S. D., Field, J. E., and Dinsmore, R. P.: Proc. Rubber Tech. Conference, London, 1938, p. 961.
(24) Gehman, S. D., and Field, J. E.: J. Applied Phys. **10**, 564 (1939).
(25) Hauser, E. A.: Naturwissenschaften **15**, 100 (1927).
(26) Hauser, E. A., Hünemörder, M., and Rosbaud, P.: Kautschuk **3**, 228 (1927).
(27) Hauser, E. A., and Mark, H.: Kolloid-Beihefte **22**, 63 (1926).
(28) Hauser, E. A., and Rosbaud, P.: Kautschuk **3**, 17 (1927).
(29) Hauser, E. A., and Smith, I. N.: India Rubber World **101**, No. 2, 31 (1939).
(30) Hauser, E. A., and v. Susich, G.: Kautschuk **7**, 145 (1931).
(31) Hengstenberg, J., and Mark, H.: Z. Krist. **69**, 271 (1928).
(32) Hintenberger, H., and Neumann, W.: Kautschuk **14**, 77 (1938).
(33) Hock, L.: Kautschuk **3**, 207 (1927).
(34) Iguchi, M., and Schossberger, E.: Kautschuk **12**, 193 (1936).
(35) Johnson, B. L., and Cameron, F. K.: Ind. Eng. Chem. **25**, 1151 (1933).
(36) Katz, J. R. Ergeb. exakt. Naturw. **4**, 161 (1925).
(37) Katz, J. R. Chem.-Ztg. **49**, 353 (1925).
(38) Katz, J. R. Naturwissenschaften **13**, 410 (1925).
(39) Katz, J. R.: Kolloid-Z. **36**, 300 (1925).

(40) Katz, J. R.: Kautschuk **5**, 6 (1929).
(41) Katz, J. R.: Trans. Faraday Soc. **32**, 77 (1936).
(42) Katz, J. R., and Bing, K.: Z. angew. Chem. **38**, 439 (1925).
(43) Katz, J. R., and Bing, K.: Gummi-Ztg. **41**, 2091 (1927).
(44) Kemp, A. R., and Peters, H.: J. Phys. Chem. **43**, 923 (1939).
(45) Long, J. D., Singer, W. E., and Davy, W. P.: Ind. Eng. Chem. **26**, 543 (1934).
(46) Lotmar, W., and Meyer, K. H.: Monatsh. **69**, 115 (1936).
(47) Mark, H., and Meyer, K. H.: Z. physik. Chem. **B38**, 395 (1938).
(48) Mark, H., and v. Susich, G.: Kolloid-Z. **46**, 11 (1928).
(49) Mark, H., and Valkó, E.: Kautschuk **6**, 210 (1930).
(50) Meyer, K. H., and Mark, H.: Ber. **61**, 1939 (1928).
(51) Meyer, K. H., and Mark, H.: *Der Aufbau der hochpolymeren organischen Naturstoffe*, p. 196. Akademische Verlagsgesellschaft, Leipzig (1930).
(52) Meyer, K. H., v. Susich, G., and Valkó, E.: Kolloid-Z. **59**, 208 (1932).
(53) Morss, H. A.: J. Am. Chem. Soc. **60**, 237 (1938).
(54) Midgley, T., and Henne, A. L.: J. Am. Chem. Soc. **59**, 706 (1937).
(55) Ott, E.: Naturwissenschaften **14**, 320 (1926).
(56) Pummerer, R., and Pahl, H.: Ber. **B60**, 2152 (1927).
(57) Pummerer, R., and v. Susich, G.: Kautschuk **7**, 117 (1931).
(58) Rossem, A. van, and Lotichius, J.: Kautschuk **5**, 2 (1929).
(59) Sauter, E.: Z. physik. Chem. **B36**, 405 (1937).
(60) Scherrer, P.: In R. Zsigmondy's *Kolloidchemie*, 4th edition, p. 409. O. Spamer, Leipzig (1922).
(61) Simard, G. L., and Warren, B. E.: J. Am. Chem. Soc. **58**, 507 (1936).
(62) Smith, W. H., and Saylor, C. P.: J. Research Natl. Bur. Standards **21**, 257 (1938).
(63) Susich, G. v.: Naturwissenschaften **18**, 915 (1930).
(64) Thiessen, P. A., and Kirsch, W.: Naturwissenschaften **26**, 387 (1938).
(65) Thiessen, P. A., and Wittstadt, W.: Z. physik. Chem. **B29**, 359 (1935).
(66) Thiessen, P. A., and Wittstadt, W.: Z. physik. Chem. **B41**, 33 (1938).
(67) Whitby, G. S.: Trans. Inst. Rubber Ind. **6**, 40 (1930).
(68) Whitby, G. S.: In Davis and Blake's *Chemistry and Technology of Rubber*. Reinhold Publishing Corporation, New York (1937).
(69) Wittstadt, W.: Kautschuk **15**, 11 (1939).

INTERATOMIC DISTANCES IN PROTEINS AND RELATED SUBSTANCES[1,2]

ROBERT B. COREY

Gates and Crellin Laboratories of Chemistry, California Institute of Technology, Pasadena, California

Received October 3, 1939

In the study of the structure of proteins the ultimate goal is a complete chemical and physical picture of the molecule,—of the nature and number of the atoms which compose it and the details of the manner of their combination. Its attainment, even in part, might be expected to bring these compounds, so fundamentally associated with life processes, under a control similar to that which the chemist now exercises over less complicated organic molecules. The great strides which have been made in this direction during recent years have resulted largely from the use of both chemical and physical techniques in the attack upon the problems of protein structure.

Chemical studies have led to the general acceptance of the idea, advanced many years ago (25), that protein molecules are made up of amino acids held together in long peptide chains by linkages between their carbon and nitrogen atoms. More than twenty amino acids have been identified as integral parts of proteins, and the development of special analytical techniques has permitted their quantitative estimation with ever-increasing accuracy. Recent results have suggested the possibility of definite periodicities in the arrangement of amino acid residues along the peptide chain (11, 12), a hypothesis which should be most fruitful in stimulating critical evaluation of existing data and in emphasizing the need of more precise analytical methods. Certainly much work remains to be done before the chemical composition of any single protein can be considered as established (34).

Preparation of increasing numbers of both plant and animal proteins in the crystalline form has encouraged the attempt to seek information concerning their structure through x-ray diffraction studies, a method which has proved so successful in determining atomic positions in crystals

[1] Contribution No. 729 from the Gates and Crellin Laboratories.

[2] Presented at the Symposium on X-ray Studies of Substances of High Molecular Weight, held at the Ninety-eighth Meeting of the American Chemical Society in Boston, Massachusetts, September, 1939, under the auspices of the Divisions of Physical and Inorganic Chemistry, Colloid Chemistry, Organic Chemistry, Paint and Varnish Chemistry, and Rubber Chemistry of the American Chemical Society.

of simpler substances. In the case of some crystalline proteins (5, 14, 20, 21) it has been possible to establish the nature and dimensions of the crystal lattice and to draw some conclusions regarding the size and shape of the molecule. However, in spite of the advances which have been made in x-ray technique, the great size and complexity of protein molecules seem to preclude the possibility of arriving at positions and relationships of individual atoms directly by these means.

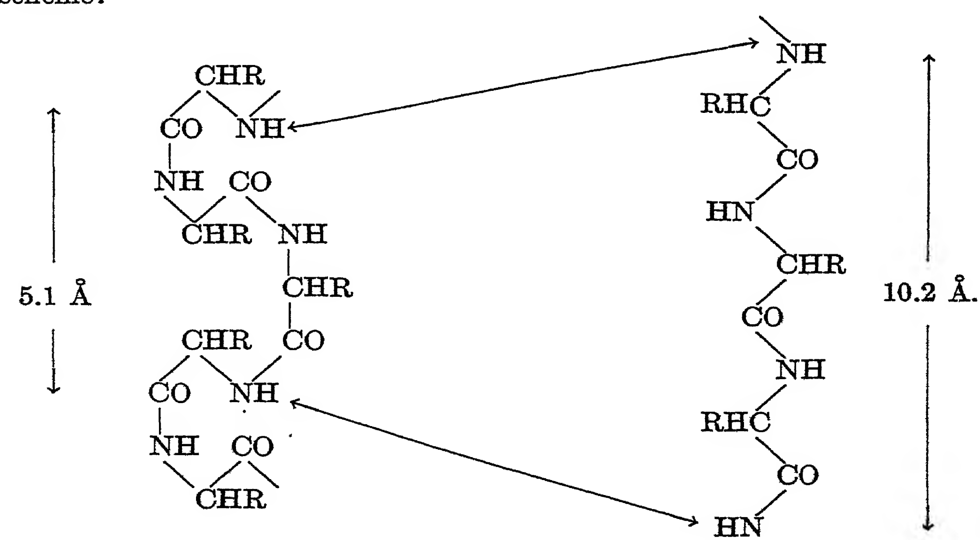
When applied to fibrous proteins, on the other hand, x-ray methods have met with considerable success in throwing light upon the intramolecular arrangements of their atoms. Silk fibroin, for example, yields excellent fiber photographs having well-defined spots and bearing marked resemblance to those obtained from comparatively simple crystalline substances. Many attempts have been made to subject these patterns to systematic analysis leading to the establishment of a fundamental unit of structure (15, 29, 30, 31, 32, 40), but it is not possible to reconcile any of the units proposed with the known facts concerning the chemical constitution of the fibroin molecule (13). All of these studies, however, indicate that the protein of silk exists as fully extended polypeptide chains arrayed nearly parallel to the axis of the fiber. Thus the photographs all reveal a definite periodicity of 7.0 ± 0.2 Å. along the fiber axis. From early crystal structure data it could be safely assumed that the distance between adjacent, chemically bonded carbon atoms is 1.54 Å. and that the corresponding separation between carbon and nitrogen is 1.3 Å. to 1.4 Å. with approximately a tetrahedral angle (109° 28′) between all bonds. In a fully extended polypeptide chain having these dimensions one amino acid residue will occupy a distance of 3.38 Å. to 3.54 Å. along the axis of the chain. These data suggest that the repetition distance, 7.0 Å. (= 2 × 3.5 Å.), found along the fiber axis of silk fibroin corresponds to two amino acid residues. Equatorial reflections, the most prominent of which represent spacings of 4.3 Å. and 4.6 A., are to be associated in some way with the packing together of adjacent chains in the fiber.

X-ray studies of other fibrous proteins lend additional significance to this picture of the fully extended polypeptide chain. In their extensive investigations of keratin, the fundamental protein constituent of animal hairs, horns, quills, etc., Astbury and his coworkers have found a characteristic spacing of 3.38 A. along the axis of the fiber when the latter is stretched to its maximum elongation (2). A corresponding periodicity of 3.3 Å. has been found in samples of stretched feather keratin (7), and a spacing along the fiber axis of 6.7 Å. (= 2 × 3.35 Å.) is reported (28) in spun filaments of blood fibroin.

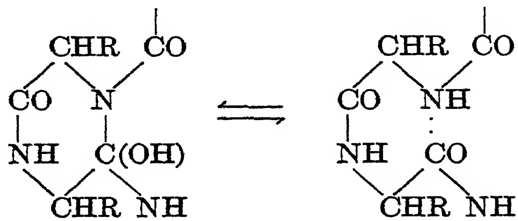
Striking uniformity is also observed in the spacings perpendicular to the fiber axis which are to be ascribed to regularity in the packing of the protein chains with their immediate neighbors. In the case of β-keratin (stretched

hair) Astbury (2) observed maxima corresponding to 4.65 Å. and 9.8 Å., the former of which he attributes to the distance between polypeptide chains held in close contact by the interaction of their =NH and =CO groups. The latter distance has been found to be practically perpendicular to the former (8) and probably lies in the direction in which the main chains are unable to approach each other more closely because of their respective side chains. Approximately these same equatorial spacings have been found in fibrin (28), and their significance is further increased by the fact that similar distances are characteristic of all denatured proteins that have been studied by x-rays (6).

Brief mention should also be made of the changes in the x-ray photographs of keratin which are observed when the specimen is stretched. Photographs of unstretched hair indicate a spacing along the fiber axis of 5.06 Å. (10). Under proper conditions hair may be stretched 100 per cent without rupture, and in this fully extended state the 5.06 Å. spacing has vanished, to be replaced by the characteristic 3.38 Å. distance associated with nearly completely extended polypeptide chains. The 9.8 Å. "side-chain" spacing perpendicular to the fiber axis is found unchanged in photographs of hair before and after stretching. The shorter (4.65 Å.) equatorial spacing, however, is found only in stretched specimens. These facts suggest that in unstretched keratin the protein is longitudinally folded in some fashion which allows of reversible extension and contraction. A mechanism for this folding was first suggested by Astbury (9) in 1930, and has been discussed by him in many subsequent publications, in which the transformation from α- to β-keratin was represented by the following scheme:



In a review published in 1935 Astbury (3) stated that "when this scheme was first put forward the precise nature of the linkage (shown dotted above) between CO and NH groups in the hexagonal folds of the α-form was left open; but we know now from accumulated x-ray crystallographic data that hexagons of the dimensions required cannot be built unless this linkage is of a covalent type that would bring the carbon and nitrogen atoms to within some $1\frac{1}{2}$ Å. of each other. A way out of the difficulty has been suggested by F. C. Frank, by postulating a sort of lactam–lactim interchange between the CO and NH groups, thus:"



More recently experimental evidence from many sources (23, 24, 26, 27, 33, 39) has shown that the existence in protein molecules of closed rings of the sort postulated is exceedingly unlikely, so that the nature of the

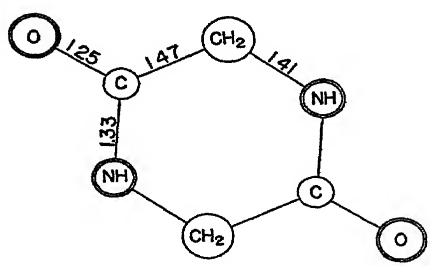


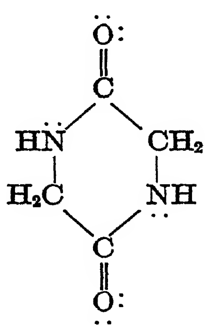
FIG. 1. A diagram of the molecule of diketopiperazine. Figures give interatomic distances in Ångström units. Angles between all bonds are close to 120°.

folding of polypeptide chains in fibrous and globular proteins is still completely undetermined. This situation has been thoroughly appreciated by Astbury who, in a paper (4) published only last year, called attention to the fact that "whether the kind of hexagonal fold postulated for the fibrous proteins provides the theme of the globular proteins too, or whether, even, it is really correct for either, it is also not yet possible to decide."

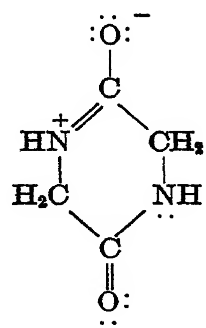
X-ray studies of fibrous proteins have been of great assistance in constructing a picture of the probable arrangement of the molecules in these substances in accord with their physical and chemical properties, and in some cases the periodicities observed are to be regarded as repetition dis-

tances within the molecules themselves. Although they thus constitute a body of experimental facts which a complete picture of the protein molecule must fully explain, they fail entirely to give any direct information regarding the distances between discrete atoms in the polypeptide chain or, indeed, in any part of the protein molecule. At present the only means of obtaining such information appears to be through x-ray investigations of crystallized products of protein hydrolysis. For this reason a series of investigations upon the crystal structures of the amino acids and polypeptides is being prosecuted in these Laboratories as a portion of a program of research upon the structure of proteins in general.

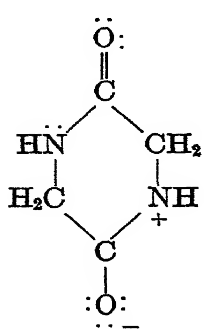
The first substance to be thus analyzed was the cyclic dipeptide diketopiperazine, or "glycine anhydride" (19). The general shape of the molecule and the interatomic distances found within it are shown diagrammatically in figure 1. The molecule is a nearly plane hexagon with the angles between all bonds 120° ± 3°. It may be expected to resonate among the structures



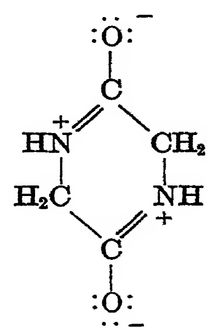
I



II



III



IV

so that the bond distances C—O and OC—N should have the values characteristic of resonance of this type. The distances found, C—O = 1.25 Å. and C—N = 1.33 Å., are in good agreement with those to be anticipated (37). Thus in urea (42) the interatomic distances within the molecule are C—O = 1.25 Å. and C—NH_2 = 1.37 Å. In thiourea (41) C—NH_2 = 1.35 Å. On the other hand, the distance N—CH_2, 1.41 Å., is surprisingly short, for there appears to be no reason for so great a departure from the normal single-bond distance, 1.47 Å., found in the compounds CH_3NO_2 (17), CH_3N_3 (36), CH_3NC (16), and $N(CH_3)_3$ (18). The C—C distance, 1.47 Å., is likewise much shorter than the normal value, 1.54 Å.

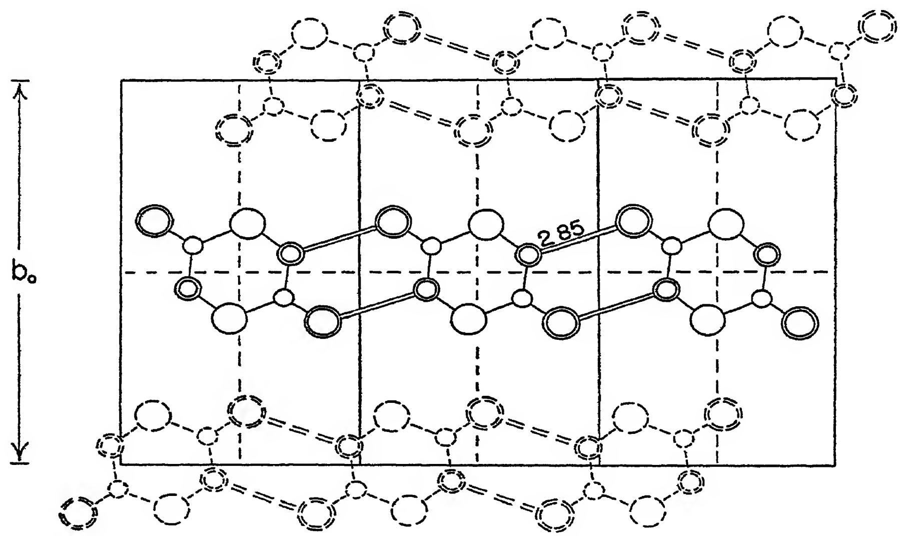


FIG. 2. A view perpendicular to the (101) plane of the crystal of diketopiperazine, showing chains of molecules held together by hydrogen bonds (double lines) 2.85 Å. in length between oxygen atoms and =NH groups.

It is hoped that work upon some substituted diketopiperazines which is now in progress in these Laboratories may throw some light upon the causes of these anomalies.

In crystals of diketopiperazine the molecules are linked together by hydrogen bonds between their respective oxygen atoms and =NH groups to form flat continuous chains throughout the structure. The positions of these chains are indicated in figure 2, which is a view of the crystal of diketopiperazine perpendicular to the plane (101). The distance between oxygen and nitrogen atoms connected by hydrogen bonds is 2.85 Å., which is in satisfactory agreement with corresponding separations in $(NH_4)H_2PO_2$, 2.81 Å. (43), and in $(NH_4)_2C_2O_4 \cdot H_2O$, 2.76 to 2.88 Å. (22).

The determination of the crystal structure of glycine (1), the simplest of the amino acids, affords additional data concerning interatomic distances within these least complicated products of protein hydrolysis. Distances and bond angles found in the glycine molecule are indicated

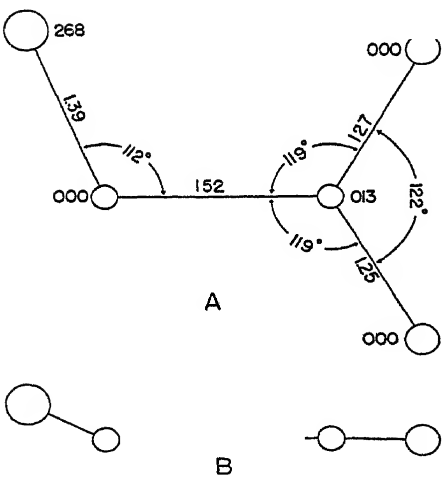


FIG. 3

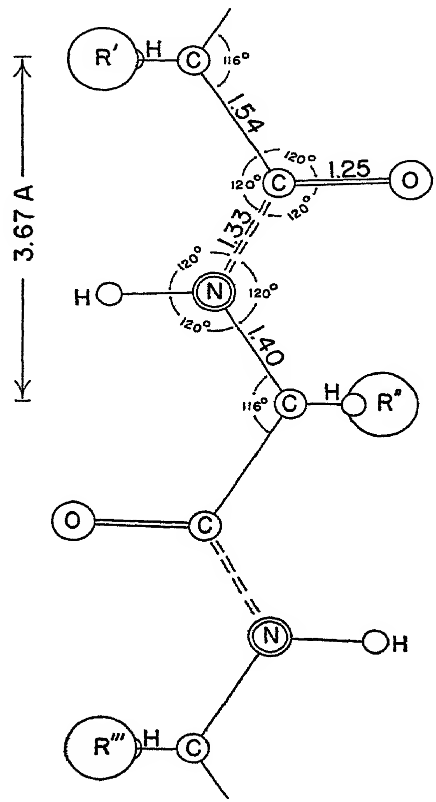


FIG. 4

FIG. 3. The molecule of glycine viewed (A) perpendicular to a plane containing the two oxygen and the α-carbon atoms, and (B) parallel to this plane.

FIG. 4. A diagrammatic representation of a fully extended polypeptide chain, based upon the interatomic distances and bond angles found in crystals of diketopiperazine and glycine. The double dashed lines represent resonating bonds (1.33 Å.) between the keto-carbon and nitrogen atoms.

in figure 3. The only value which departs from normal is the C—N separation, 1.39 Å., which, however, is in agreement with the analogous distance found in diketopiperazine. Interatomic distances found within the molecules of these two compounds are summarized in table 1.

Even from these limited data it should be possible to make some definite

statements regarding the linkages present in the polypeptide chain in proteins. The distance between the nitrogen atom and the α-carbon atom is 1.40 Å., with a probable error of 0.02 Å., in both diketopiperazine and glycine. Although this distance is about 0.07 Å. shorter than that to be expected from the normal covalent radii of the atoms (35, 38), there seems to be little doubt that this new distance found in these hydrolytic products of proteins exists also in the protein molecule itself. Whether the abnormally short C—C distance found in the cyclic dipeptide prevails also in the long peptide chains of proteins is a question which can be settled only by complete crystal structure studies of simple straight-chain polypeptides and substituted diketopiperazines. There seems to be every reason for believing that the type of resonance evidenced by diketopiperazine, urea, etc., will be present in the polypeptide chain so that the distance between the nitrogen atom and its adjacent keto-carbon atom must be very near

TABLE 1

Corresponding interatomic distances found in molecules of diketopiperazine and glycine

	DIKETOPIPERAZINE	GLYCINE
	Å.	*Å.*
C—O	1.25	1.26
α-C—N	1.41	1.39
C—C	1.47	1.52
OC—N	1.33	

to 1.33 Å. For the same reason the oxygen atom must be about 1.25 Å. from the keto-carbon atom.

A diagram of a fully extended polypeptide chain in which these dimensions have been incorporated is shown in figure 4. The bond angles about the carbonyl carbon and the nitrogen atoms are assumed to be 120°, as found in diketopiperazine. Although the true angles between these bonds doubtless differ somewhat from this value, it is unlikely that any of them departs from it by more than 5°. Around the α-carbon atom the bonds are assumed to be arranged in nearly tetrahedral fashion, except that the C—C—N bond angle is taken as 116°, a mean of the values found in glycine (112°) and diketopiperazine (120°). The fully extended chain is coplanar, with the carbonyl oxygen atoms and amide hydrogen atoms included in the plane. The distance along the chain corresponding to one amino acid residue is 3.67 Å., which is somewhat greater than that estimated from x-ray photographs of any fibrous protein. This fact suggests that in these substances the chain is never fully extended in a truly coplanar configuration but that interactions, steric or otherwise, with its immediate neighbors cause slight distortions, probably involving rotations about the C—C

bond. This view finds some confirmation if the attempt is made to place these fully extended chains side by side at about 4.65 Å. from each other and in such a fashion as to form hydrogen bonds between the =CO and =NH groups of adjacent chains. It is then found that all such arrangements result in steric interferences between side chains, or between side chains and oxygen atoms, as long as complete coplanarity of the polypeptide chain is rigidly maintained. However, for certain relative positions it is possible to avoid these conflicts by slight rotations back and forth about the carbon–carbon bond. Unfortunately, existing x-ray data are inadequate as reliable guides in making a selection among the many possible configurations which present themselves.

Thus, although our present information is insufficient to supply a definite solution of even the simpler problems of the atomic arrangements in proteins, there seems to be ample justification for the belief that more precise knowledge of the interatomic distances within the polypeptide chain, together with the results of more refined x-ray and chemical techniques, will provide experimental foundation for the development of a critical theory of protein structure.

It is a pleasure to acknowledge the kindness of Professor Linus Pauling, whose helpful discussions have been a great aid in the preparation of this paper.

REFERENCES

(1) Albrecht, G., and Corey, R. B.: J. Am. Chem. Soc. **61**, 1087 (1939).
(2) Astbury, W. T.: Trans. Faraday Soc. **29**, 193 (1933).
(3) Astbury, W. T.: J. Textile Inst. **27**, 282 (1936).
(4) Astbury, W. T.: Trans. Faraday Soc. **34**, 378 (1938).
(5) Astbury, W. T., and Lomax, R.: Nature **133**, 795 (1934).
(6) Astbury, W. T., and Lomax, R.: J. Chem. Soc. **1935**, 846.
(7) Astbury, W. T., and Marwick, T.: Nature **130**, 309 (1932).
(8) Astbury, W. T., and Sisson, W. A.: Proc. Roy. Soc. (London) **A150**, 533 (1935).
(9) Astbury, W. T., and Woods, H. J.: Nature **126**, 913 (1930).
(10) Astbury, W. T., and Woods, H. J.: Phil. Trans. **A232**, 333 (1933).
(11) Bergmann, M.: Harvey Lectures **31**, 37 (1936); Chem. Rev. **22**, 423 (1938).
(12) Bergmann, M., and Niemann, C.: J. Biol. Chem. **115**, 77 (1936); **118**, 301 (1937); Science **86**, 187 (1937).
(13) Bergmann, M., and Niemann, C.: J. Biol. Chem. **122**, 577 (1938).
(14) Bernal, J. D., and Crowfoot, D.: Nature **133**, 794 (1934).
(15) Brill, R.: Ann. **434**, 204 (1923).
(16) Brockway, L. O.: J. Am. Chem. Soc. **58**, 2516 (1936).
(17) Brockway, L. O., Beach, J. Y., and Pauling, L.: J. Am. Chem. Soc. **57**, 2693 (1935).
(18) Brockway, L. O., and Jenkins, H. O.: J. Am. Chem. Soc. **58**, 2036 (1936).
(19) Corey, R. B.: J. Am. Chem. Soc. **60**, 1598 (1938).
(20) Crowfoot, D.: Nature **136**, 591 (1935).

(21) Crowfoot, D.: Proc. Roy. Soc. (London) **A164**, 580 (1938).
(22) Hendricks, S. B., and Jefferson, M. E.: J. Chem. Phys. **4**, 102 (1936).
(23) Haurowitz, T.: Z. physiol. Chem. **256**, 28 (1938).
(24) Haurowitz, T., and Astrup, T.: Nature **143**, 118 (1939).
(25) Hofmeister, F.: Ergeb. Physiol. biol. Chem. exptl. Pharmakol. **1**, 759 (1902).
(26) Huggins, M. L.: J. Am. Chem. Soc. **61**, 755 (1939).
(27) Jenkins, G. I., and Taylor, T. W. J.: J. Chem. Soc. **1937**, 495.
(28) Katz, J. R., and de Rooy, A.: Rec. trav. chim. **52**, 742 (1933).
(29) Kratky, O.: Z. physik. Chem. **B5**, 297 (1929).
(30) Kratky, O., and Kuriyama, S.: Z. physik. Chem. **B11**, 363 (1931).
(31) Meyer, K. H., and Mark, H.: *Der Aufbau der hochpolymeren organischen Naturstoffe*, p. 222. Akademische Verlagsgesellschaft, Leipzig (1930).
(32) Meyer, K. H., and Mark, H.: Ber. **61**, 1932 (1928).
(33) Neurath, H., and Bull, H. D.: Chem. Rev. **23**, 427 (1938).
(34) Niemann, C.: Cold Spring Harbor Symposia Quant. Biol. **6**, 58 (1938).
(35) Pauling, L.: *The Nature of the Chemical Bond*, p. 154. Cornell University Press, Ithaca, New York (1939).
(36) Pauling, L., and Brockway, L. O.: J. Am. Chem. Soc. **59**, 13 (1937).
(37) Pauling, L., and Brockway, L. O.: J. Am. Chem. Soc. **59**, 1223 (1937).
(38) Pauling, L., and Huggins, M. L.: Z. Krist. **87**, 205 (1934).
(39) Pauling, L., and Niemann, C.: J. Am. Chem. Soc. **61**, 1860 (1939).
(40) Trogus, C., and Hess, K.: Biochem. Z. **260**, 376 (1933).
(41) Wyckoff, R. W. G., and Corey, R. B.: Z. Krist. **81**, 386 (1932).
(42) Wyckoff, R. W. G., and Corey, R. B.: Z. Krist. **89**, 462 (1934).
(43) Zachariasen, W. H., and Mooney, R. C. L.: J. Chem. Phys. **2**, 34 (1934).

X-RAY DIFFRACTION STUDY OF THE STRUCTURE OF GLASS[1]

B. E. WARREN

George Eastman Laboratory of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts

Received November 24, 1939

CONTENTS

I. INTRODUCTION

Glass is usually called an undercooled liquid, the name suggesting that, although it has many of the mechanical properties of a true solid, it differs from the crystalline form of matter by not having passed through a sharp or definite transition in solidifying from the melt. From the x-ray studies we shall conclude that glass is similar to a liquid in that both are amorphous forms of matter. In one respect, however, their structures differ: in a glass each atom has permanent neighbors at a fairly definite distance, while in a liquid the neighbors about any atom or molecule are continually changing.

The x-ray diffraction pattern of a glass consists of one or more broad diffuse rings. It is distinctly different from the powder pattern of a crystalline material, which shows a large number of fairly sharp rings. Most of the early attempts to analyze amorphous patterns approached the problem from the point of view of crystalline diffraction, and tried to explain the diffraction bands as Bragg reflections from layers of atoms such as the planar layers in crystalline structures. Only recently have x-ray workers realized that an x-ray scattering pattern showing maxima

[1] Presented at the Symposium on X-ray Studies of Substances of High Molecular Weight, held at the Ninety-eighth Meeting of the American Chemical Society at Boston, Massachusetts, September, 1939, under the auspices of the Divisions of Physical and Inorganic Chemistry, Colloid Chemistry, Organic Chemistry, Paint and Varnish Chemistry, and Rubber Chemistry of the American Chemical Society.

and minima can be produced solely by the existence of a few fairly definite interatomic distances in the sample material.

With the development of the Fourier method of analysis, it is now possible to carry through a straightforward and rigorous analysis of the x-ray pattern of a glass. From this type of analysis one gets directly the interatomic distances in the glass and the number of neighboring atoms about each kind of atom. From these results certain important conclusions can be drawn, and for the simple glasses the results are often sufficient to establish a fairly definite picture of the structure.

A wide variety of materials can be brought into the glassy state. Among the elements selenium is the well-known example; in the inorganic field the silicates, borates, and phosphates are ready glass formers, and in the field of organic chemistry glasses are formed by many substances, particularly by molecules containing a fair number of hydroxyl radicals. The diffraction patterns of the inorganic glasses are the simplest to interpret;

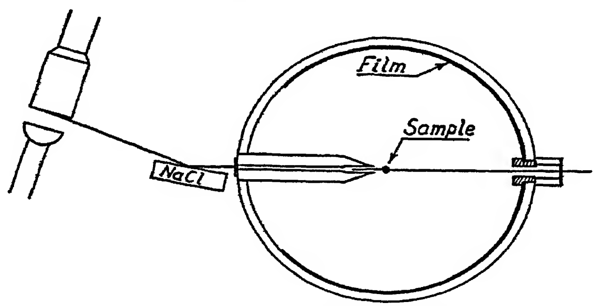


FIG. 1. Camera for x-ray diffraction patterns of glass

furthermore, these are the glasses of practical interest. Accordingly, the discussion in this paper will be confined to silicate and borate glasses.

II. EXPERIMENTAL METHODS

X-ray diffraction patterns of glass are usually made in a cylindrical camera of about 5 cm. radius. The schematic arrangement is shown in figure 1. A highly monochromatic beam is obtained by allowing the radiation from a molybdenum or copper target tube to fall upon a rock salt crystal set at the correct angle to reflect the K_α line. The monochromatic reflected beam passes through a collimating system and falls upon the glass sample, which is usually in the form of a rod about 0.5 to 1.0 mm. in diameter. The diffracted radiation is recorded on a photographic film placed inside the cylindrical camera. To eliminate the scattering of the main beam by air, the camera is evacuated during the exposure. Figures 2 and 3 are typical x-ray diffraction patterns of glass.

The diffraction patterns are microphotometered and the microphotometer records changed to intensity curves in the usual way. The final experimental result is a curve giving the intensity of x-ray scattering (in arbitrary units) as a function of the angle of scattering. Intensity curves for vitreous silica and for a series of soda–silica compositions are shown in figures 4 and 9.

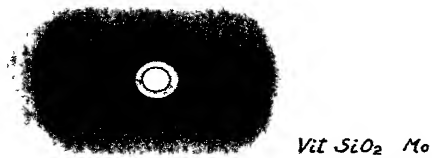


FIG. 2 X-ray pattern of vitreous silica taken with Mo K_α radiation (λ = 0.710 Å.)

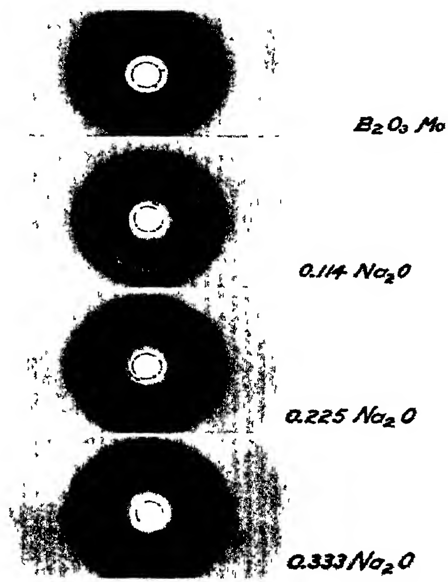


FIG. 3. X-ray patterns of soda–boric oxide glass taken with Mo K_α radiation (λ = 0.710 Å.)

III. THEORY OF FOURIER ANALYSIS

From a typical glass diffraction pattern, containing three or four diffuse rings, the only quantity which can be determined directly and uniquely is the radial distribution function. This is obtained from a Fourier analysis of the experimental x-ray scattering curve, and gives directly the average number of atoms to be found at any distance from a given atom. The method of Fourier analysis is equally applicable to a liquid,

a glass, or a powdered crystalline material, so that in using this method one is not making any *a priori* assumptions as to the structure of the material in question.

For simplicity, the theory of the method of Fourier analysis will be developed here for the simple case of a material containing only one kind of atom. A monochromatic x-ray beam falls upon an array of atoms, and the radiation scattered at an angle 2θ to the direction of the primary beam is observed at a point P. The amplitude of the unmodified scattered radiation is given in electron units (using the radiation scattered by a single electron according to classical theory as the unit) by equation 1.

$$E_p = \sum_n f_n \, e^{\frac{2\pi i}{\lambda}(\bar{S}-\bar{S}_0)\cdot\bar{r}_n} \tag{1}$$

Σ indicates summation over all the atoms in the sample, $\bar{r}_n$ is a vector giving the position of atom n, f_n is the atomic scattering factor of atom n, and $\bar{S}$ and $\bar{S}_0$ are unit vectors giving the direction of the scattered and primary radiation.

For a crystalline material the summation of equation 1 can be carried out, and an expression secured for the resultant amplitude and intensity. This intensity expression will show sharp maxima in the usual directions given by the Bragg law. If one is not at liberty to assume the material to be crystalline, it is not possible to carry out the summation indicated in equation 1. Multiplying equation 1 by its conjugate complex quantity, and then averaging the resulting expression for intensity, as the rigid array of atoms is allowed to take all possible orientations in space, one obtains the Debye equation (4):

$$I = \sum_m \sum_n \frac{f_m f_n \sin Sr_{mn}}{Sr_{mn}} \tag{2}$$

$$S = \frac{4\pi \sin\theta}{\lambda}$$

$$r_{mn} = \text{distance from atom } m \text{ to atom } n$$

In applying equation 2 to a glass, it is obviously not necessary actually to rotate the sample, since any configurations which may exist will be found many times with all possible orientations. For a material consisting of only one kind of atom, and with the assumption that on the average each atom is surrounded in the same way as every other atom, equation 2 becomes

$$I = Nf^2 \sum_n \frac{\sin sr_n}{sr_n} \tag{3}$$

where N is the effective number of atoms in the sample.

Introducing a radial distribution function such that $4\pi r^2\rho(r)\,dr$ is the number of atoms between distances r and $r + dr$ from any atom, equation 3 becomes

$$I = Nf^2\left\{1 + \int_0^\infty 4\pi r^2\rho(r)\;\frac{\sin sr}{sr}\,dr\right\} \tag{4}$$

Rearranging this gives

$$s\left(\frac{I}{Nf^2} - 1\right) = 4\pi\int_0^\infty r(\rho - \rho_0)\sin sr\,dr + 4\pi\int_0^\infty r\rho_0 \sin sr\,dr \tag{5}$$

Except for very small values of s, the second integral is zero. Inverting the rest of the equation by the Fourier integral theorem, we obtain finally

$$4\pi r^2\rho(r) = 4\pi r^2\rho_0 + \frac{2r}{\pi}\int_0^\infty si(s)\sin rs\,ds \tag{6}$$

$$i(s) = \left(\frac{I}{Nf^2} - 1\right)$$

ρ_0 = average density of the sample, in atoms per unit volume

The quantity $si(s)$ is obtained directly from the experimental scattering curve. The integration involved in equation 6 is then carried out for a number of different values of r. The resulting values of $4\pi r^2\rho(r)$ plotted against r give the radial distribution of atoms about any average atom. Equation 6 was first derived by Zernike and Prins (14), and was first applied by Debye and Menke (5) in an x-ray study of liquid mercury.

For a material containing more than one kind of atom, such as vitreous silica, equation 6 takes the more general form (12):

$$\sum K_m 4\pi r^2\rho_m = \sum K_m 4\pi r^2\rho_0 + \frac{2r}{\pi}\int_0^\infty si(s)\sin rs\,ds \tag{7}$$

Σ indicates summation over an assumed unit of composition, K_m is the effective number of electrons in atom m, ρ_m is the electron density surrounding atom m, and ρ_0 is the average number of electrons per unit volume.

The weighted radial distribution function $\Sigma K_m 4\pi r^2\rho_m$ given by equation 7 is obtained from the experimental x-ray scattering curve. The result is unique, in that no assumptions as to structure or crystallinity are involved. The significance of the distribution function is best explained in terms of the specific examples which follow.

IV. FOURIER ANALYSIS OF VITREOUS SILICA

A simple stable glass, and one that is very convenient for x-ray study, is vitreous silica (fused quartz). For the details of the x-ray analysis of

vitreous silica, reference should be made to the original paper (12). The diffraction patterns were made in an evacuated camera of radius 4.40 cm., using Cu K_α and Mo K_α radiation monochromated by reflection from a rock salt crystal. Intensity curves were obtained from the microphotometer records of the diffraction patterns. The intensity curves were put upon an absolute basis (electron units per SiO_2) by consideration of the fact that at large values of sin θ/λ the experimental curve must approach the curve for independent scattering by the atoms. The experimental curve must be corrected for absorption, polarization, and Compton modi-

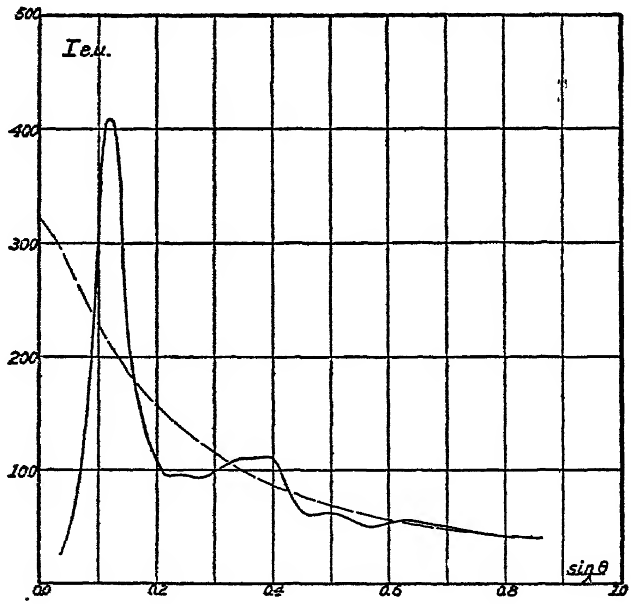


FIG. 4. X-ray intensity curve for vitreous silica in electron units per SiO_2. Dashed line, independent scattering per SiO_2.

fied scattering. The final intensity curve for vitreous silica is shown in figure 4.

From the intensity curve the quantity $si(s)$ is calculated and plotted as a function of s. The integration involved in equation 7 is then carried out for about forty different values of r ranging from 0 to 8 Å. The integration can be done either graphically or with a harmonic analyzer. The result is a series of values of $\Sigma K_m 4\pi r^2 \rho_m$. These values plotted against r give the radial distribution curve for vitreous silica shown in figure 5.

Since there are two kinds of atoms in silica, the curve of figure 5 is really two distribution curves superimposed,—the distribution of neigh-

boring atoms about a silicon atom, and the distribution about an oxygen atom. The positions of the peaks give the distances of neighboring atoms from a silicon or an oxygen atom. From the area under a peak, it is generally possible to calculate the number of neighbors at that distance. The significance of the different peaks is established by comparing the interatomic distances given by the radial distribution curve with the various known interatomic distances determined from crystalline structures.

The first peak of figure 5 occurs at about 1.62 Å. Since this is the silicon–oxygen distance which has been found in all crystalline forms of silica and in crystalline silicates, the peak is identified as representing the

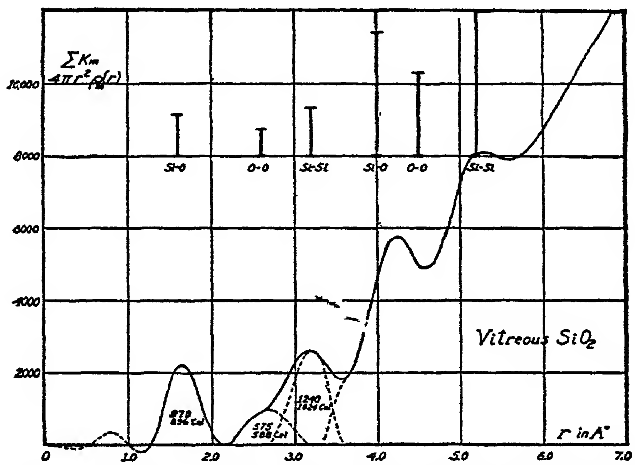


FIG. 5. Radial distribution curve for vitreous silica

silicon–oxygen distance in the glass. From the area under the first peak we can calculate the number of oxygens about each silicon. If there are n oxygens bonded to each silicon, there must be $n/2$ silicons bonded to each oxygen. From the atomic scattering factors for silicon and oxygen, it is found that the best average value for the effective number of electrons per atom is for silicon $K = 16.0$ and for oxygen $K = 7.0$. The area under the first peak will be

$$A = 1 \times 16 \times n \times 7.0 + 2 \times 7.0 \times (n/2) \times 16$$

Equating this to the observed peak area 970 elect^2, and solving for n:

$$n = \frac{970}{1 \times 16.0 \times 7.0 \times 2} = 4.3$$

[2] For a discussion of early work bearing on this question see reference 6.

Since the silicon atom is always found tetrahedrally surrounded by four oxygens at a distance of about 1.62 Å. in all crystalline silicates, the measured value 4.3 is interpreted as meaning 4 within the limits of experimental error. If each silicon is tetrahedrally surrounded by four oxygens, then from the composition of the glass each oxygen must be bonded to two silicons. Assuming that the four oxygens about a silicon are tetrahedrally arranged, the oxygen–oxygen distance should be

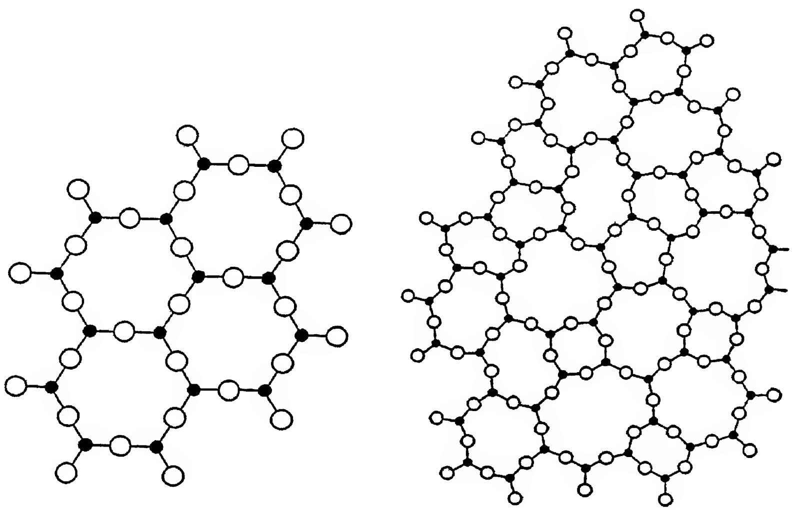
$$\text{O—O} = 1.62 \sqrt{8/3} = 2.65 \text{ Å.}$$

The second and third peaks are not well resolved, but the existence of a peak at about 2.65 Å., of about the right area to correspond to six oxygen neighbors about each oxygen, is clearly indicated on the distribution curve.

If the two bonds to an oxygen are roughly diametrically opposite, the silicon–silicon distance will be approximately twice the silicon–oxygen distance, or about 3.2 Å. A distinct peak at 3.2 Å. shows that this is the case. The first three peaks indicate that each silicon is tetrahedrally surrounded by four oxygens at a distance of 1.62 Å. and each oxygen is bonded to two silicons, the two bonds being roughly diametrically opposite. So far as concerns the satisfying of these requirements, the relative orientation of one tetrahedral group with respect to a neighboring group, around the direction of the connecting Si—O—Si bond, can be practically random. In spite of this randomness, a number of definite interatomic distances exist. These distances are indicated on figure 5 by a series of vertical lines, the heights of the lines being proportional to the expected peak areas. The oxygen–second oxygen distance depends upon the relative orientation of two tetrahedral groups about their connecting Si—O—Si bond, and the value 4.5 Å. which is given represents only an average distance. The peak at 4.2 Å. in the experimental curve is interpreted as an unresolved average of the Si—O peak at 4.0 Å. and the average O—O separation at 4.5 Å. For distances greater than the silicon–second silicon distance of 5.2 Å., the various interatomic distances depend upon the relative orientation of the tetrahedral groups, and the distribution curve rapidly smooths out.

It should be pointed out that the distribution curve of figure 5 is a unique result for vitreous silica, since the application of the Fourier method of analysis involves no assumption as to whether the material is truly amorphous or crystalline. The x-ray results are completely explained by picturing glassy silica as a random network in which each silicon is tetrahedrally surrounded by four oxygens and each oxygen is bonded to two silicons, the two bonds to an oxygen being roughly diametrically opposite. The orientation of one tetrahedral group with respect to a neighboring group about the connecting Si—O—Si bond can be practically

random. This is the simplest picture of silica glass, free from all specialized assumptions, which will completely explain the x-ray diffraction pattern. There is a definite scheme of structure involved: each atom has a definite number of nearest neighbors at a definite distance, but no unit of structure repeats itself identically at regular intervals in three dimensions, and hence the material is not crystalline. This is essentially the picture of an oxide glass which Zachariasen (13) deduced from consideration of the chemical composition. Figure 6, taken from Zachariasen's paper, illustrates schematically in two dimensions the irregular structure of a glass as distinguished from the regularly repeating structure of a crystal



a b

FIG. 6. Schematic representation in two dimensions of the difference between a crystal and a glass: a, crystal; b, glass. (W. H. Zachariasen.)

It should be emphasized that there is no such thing as an SiO_2 molecule in silica glass. Instead, the glass is a high-polymer random network in which each silicon is bonded to four oxygens and each oxygen is bonded to two silicons. The non-existence of an SiO_2 molecule in silica glass should not be surprising, since the extensive x-ray studies of crystalline silicates have shown that the fundamental law of structural silicate chemistry is the tetrahedral bonding of four oxygens about each silicon, and that molecules have no existence in silicate structures.

V. THE CRYSTALLITE THEORY OF GLASS

While the Fourier analysis gives uniquely the coördination scheme, that is, the number of neighbors and their distances, it does not answer the

specific question as to whether or not the material can be considered crystalline. A number of workers have interpreted the x-ray diffraction patterns of vitreous silica as due to extremely small crystals of the high-temperature crystalline form of silica, cristobalite (6, 8).

Figure 7 shows the diffraction patterns of vitreous silica and cristobalite taken under identical conditions. Figure 8 shows the microphotometer records of these patterns. The strongest ring in the cristobalite pattern comes at very nearly the position of the first broad peak in the vitreous silica pattern. Since the breadth of the lines on a crystal powder pattern increases continuously with decreasing size of crystalline particles, it is evident that if one postulates in vitreous silica cristobalite crystals sufficiently small, the strong peak of cristobalite would broaden out to give

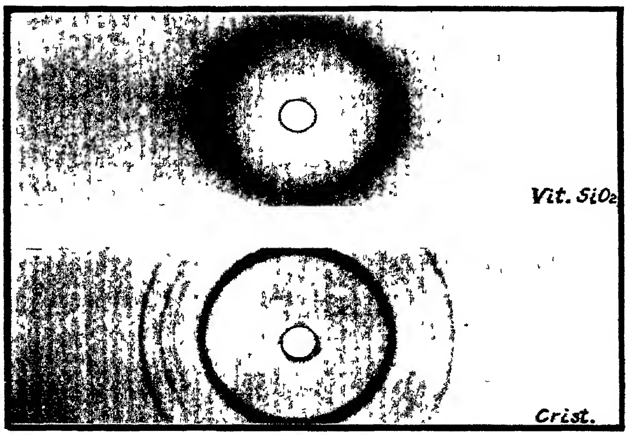


FIG. 7. X-ray diffraction patterns: above, vitreous silica, below, cristobalite

the appearance of the vitreous pattern. From a formal point of view this interpretation is perfectly sound, and since cristobalite is the stable form of silica through the temperature range in which glassy silica stiffens up, it is not at all unreasonable to expect cristobalite crystals in the glass.

If we assume tentatively that the x-ray pattern of vitreous silica is due to small cristobalite crystals, the size of these assumed crystallites can be calculated from the line breadth. Changing the microphotometer record to an intensity curve, and measuring the half-intensity breadth of the strong peak, the average particle dimension is calculated by the Bragg particle size equation (3)

$$L = \frac{0.89\lambda}{B \cos \theta} \quad \frac{0.89 \times 1.54}{0.181 \times 0.98} = 7.7 \text{ Å} \tag{8}$$

The edge of the unit cell of cristobalite is about 7.0 Å., and the crystallite hypothesis has therefore forced us to postulate cristobalite crystals scarcely larger than one unit cell in order to explain the observed peak width.

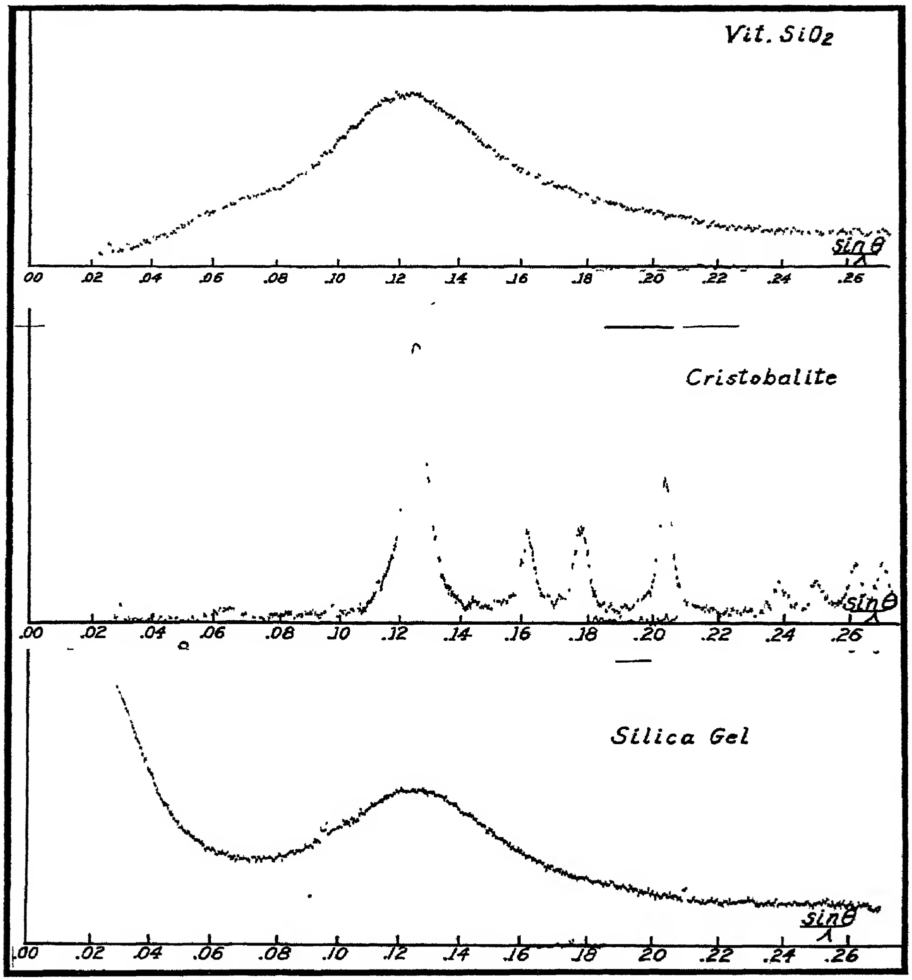


FIG. 8. Microphotometer records of x-ray diffraction patterns: a, vitreous silica; b, cristobalite; c, dried commercial silica gel.

The essence of the idea of crystallinity is regular repetition, and in a particle containing only one unit cell repetition does not exist. There is then a very real objection to the cristobalite crystallite description of silica glass (9, 10). In order to conform to the experimental facts, it

requires the extension of the term "crystal" to blocks of matter so small that the term ceases to have any meaning.

Another question of interest concerns the homogeneity of glass. Is glass an aggregation of small groups of atoms with breaks and voids between these groups, or is it a homogeneous medium with a continuous scheme of bonding? In other words, is glass like a pail full of pebbles or a pail full of water? This question can be answered by comparing the small angle scattering of silica glass and silica gel shown on figure 8. The main peak is the same for both the silica glass and the silica gel, but whereas for the glass the intensity falls to zero at small scattering angles, for the gel it rises to high values.

Small angle scattering is due to large scale inhomogeneities in the sample (large with respect to atomic dimensions). The diffraction patterns indicate that the dried silica gel comprises silica particles having about the same structure as the glass, but with gaps and voids between these particles of the order of 10 to 100 Å. In marked contrast to the gel, the pattern of silica glass shows no small angle scattering, indicating a completely homogeneous medium without discrete particles, gaps, or voids. This means that the random tetrahedral network is one in which there is a continuous scheme of bonding.

The x-ray data for vitreous silica are completely explained on the basis of the random tetrahedral network picture of the glass. However, x-ray analysis does not prove that there is no crystalline cristobalite in glassy silica. Here and there one would expect to find the cristobalite configuration existing for a short distance, as one of a continuous variety of configurations. The x-ray analysis does, however, prove that the major part of the material cannot be in the form of cristobalite crystals of sufficient size to give the term any meaning. The devitrification of vitreous silica to cristobalite, over a wide range of temperatures, is probably facilitated by the presence of a small amount of crystalline cristobalite acting as seed material. Trostel (7) has found that a silica glass containing a small trace of quartz can be devitrified at about 1300°C. to give mainly quartz as the devitrification product.

The scheme of coördination about each atom is the same in silica glass and in crystalline cristobalite. If a man sitting on a silicon atom could look no farther than the nearest and next nearest atoms, he would not know whether he was in a piece of silica glass or in a cristobalite crystal. The glass is a form of matter in which the coördination scheme is the same as in the crystalline phase, but which cooled too rapidly from a viscous melt to allow any disentangling and subsequent rebuilding of an orderly structure.

VI. STRUCTURE OF SODA–SILICA GLASS

Soda–silica glass can be formed with compositions ranging from pure silica to the equimolal composition $Na_2O \cdot SiO_2$. The soda–silica glasses form an interesting binary system of great practical importance. X-ray intensity curves for a series of compositions are given by figure 9. It should be noted that the change in the intensity curves is a perfectly continuous change with increase in soda content. From a Fourier analysis

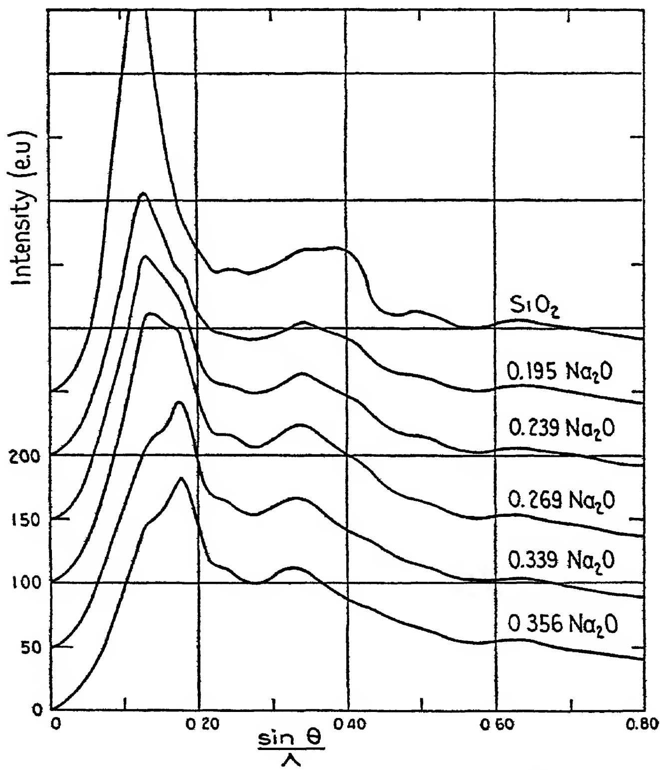


FIG. 9. X-ray intensity curves for soda–silica glass

of the intensity curves, the radial distribution curves of figure 10 are obtained (11).

For each composition the first peak occurs at about 1.62 Å., corresponding to the well-known silicon–oxygen distance in silicates. The number of oxygens around a silicon can be calculated from the area under the peak. The values obtained are shown in figure 10. Although the values run a little above 4, it is assumed that they should be interpreted as 4 within the limits of experimental error. Of course it might be argued that a

value of 4.5 could be interpreted as indicating part of the silicons surrounded by four oxygens and part by five oxygens. However, this idea is dismissed as completely improbable, in view of the well-established tetrahedral bonding of four oxygens about each silicon in all silicates so far studied.

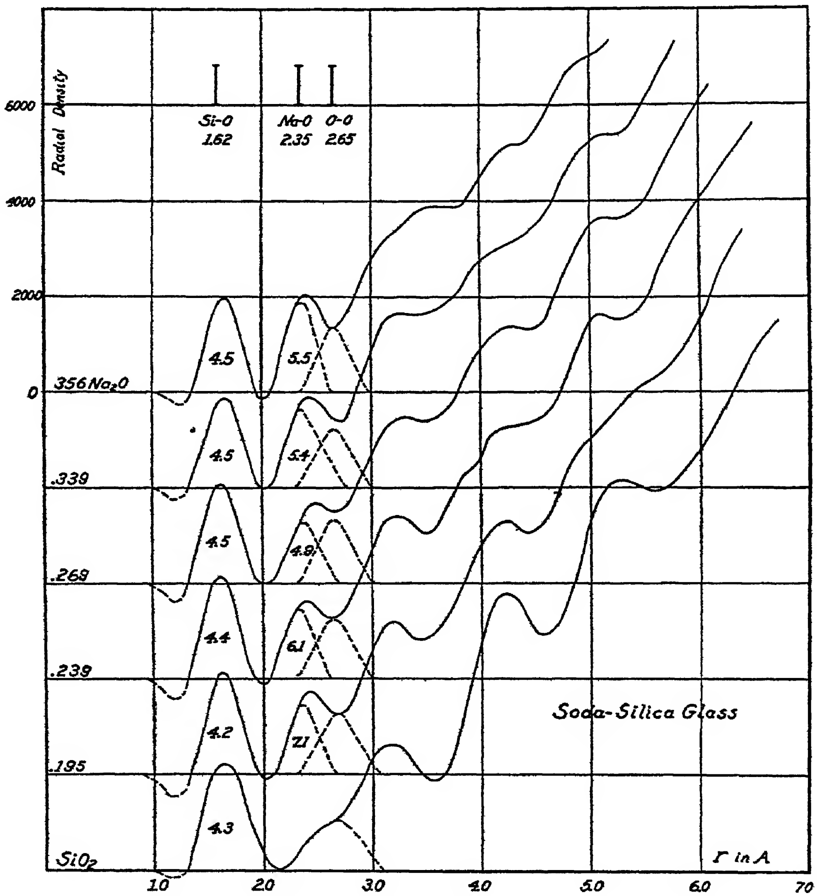


FIG. 10. Radial distribution curves for soda-silica glass

The next two distances, Na—O and O—O, are too close together to give resolved peaks. From the known tetrahedral bonding and the composition of the glass, the area of the O—O peak at 2.65 Å. can be calculated. These peaks, which are shown dotted on figure 10, are subtracted from the distribution curve, leaving a peak at about 2.35 Å. These residual peaks are interpreted as representing the Na—O separation, since

the distance 2.35 Å. corresponds to the sum of the atomic radii for Na^+ and O^{--}. From the area of the residual peaks the number of oxygens about a sodium atom is readily calculated. The values listed on figure 10 have an average of about 6, which is a reasonable coördination number for oxygen atoms about a sodium atom.

A soda–silica glass is not a structure comprising silica molecules, SiO_2, and soda molecules, Na_2O. The sodium exists as the ion Na^+, and all of the oxygen in the composition takes part in forming the tetrahedral coördination of four oxygens about each silicon. Since the total number of oxygens is greater than twice the number of silicons, it is not possible

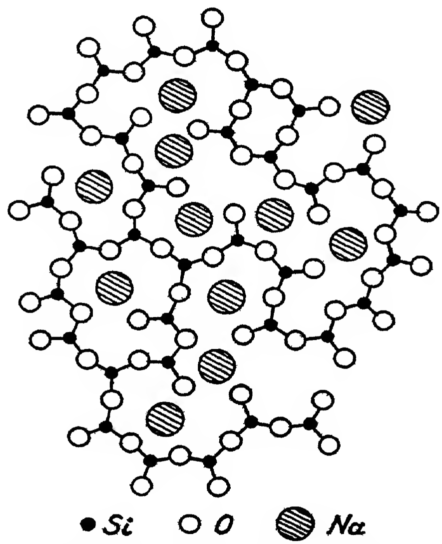


FIG. 11. Schematic representation in two dimensions of the structure of soda–silica glass.

for each oxygen to be bonded to two silicons. Part of the oxygens must be bonded to two silicons and part bonded to only one silicon.

Figure 11 represents schematically in two dimensions the structure of soda–silica glass as deduced from the present x-ray study. Because the real structure exists in three dimensions, it is necessary to take certain liberties in making a schematic two-dimensional representation. In three dimensions, each silicon is tetrahedrally surrounded by four oxygens, and in the two-dimensional representation each silicon is shown surrounded only by three oxygens. The oxygens are correctly represented, some of them bonded between two silicons, and others bonded only to one silicon. The sodium ions, Na^+, are shown in various holes in the irregular silicon–

oxygen network. This figure represents very well the essential scheme of structure in a soda–silica glass. There is a definite scheme of co-ordination,—each silicon tetrahedrally surrounded by four oxygens and part of the oxygens bonded between two silicons and part only to one silicon. The sodium ions are held rather loosely in the various holes in the silicon–oxygen network and are surrounded on the average by about six oxygens. Although it is a perfectly definite scheme of structure, there is no regular repetition in the pattern and the structure is non-crystalline.

Because there is no regular repetition, it is evident why soda–silica glass has no definite chemical composition. Starting with vitreous silica, there is a random network in which each silicon is tetrahedrally surrounded by four oxygens and each oxygen is bonded between two silicons. As the soda content is increased, the number of oxygens becomes greater than twice the number of silicons, and hence more and more of the oxygens are bonded only to one silicon. The sodium ions take the best places they can find in the various openings in the silicon–oxygen network.

Electrical conduction is readily understood in terms of figure 11. Owing to the temperature motion, the atoms of the silicon–oxygen network are oscillating about mean positions. Under the influence of the electric field, the loosely bound sodium ion falls step-fashion from one network hole to another as conditions make the step possible.

The lowering of the softening temperature of silica glass by addition of soda is also understood from figure 11. In vitreous silica there is a complete system of bonding in three dimensions, each silicon to four oxygens and each oxygen to two silicons. With the addition of soda, breaks occur in the silicon–oxygen framework, owing to some of the oxygens being bonded only to one silicon. With more and more breaks in the strong silicon–oxygen network, the structure is no longer rigidly braced in three dimensions. Displacements and changes in configuration in the silicon–oxygen network can take place at temperatures which are still too low for disrupting any appreciable number of silicon–oxygen bonds.

Because there is no scheme of repetition in the glass, no two points are exactly identical. There are points with widely varying degrees of weakness, at which flow or breaking can occur at a continuous variety of temperatures. It is readily understood, therefore, why glass gradually softens, rather than having a definite melting point as a crystal does.

Vitreous silica has an extremely small coefficient of thermal expansion and is a very stable glass. For many purposes it would be the ideal glass to use, if it were not for the fact that the temperatures required to make and work it are inconveniently high. Soda is added to silica merely to

soften the glass and to bring the working temperature down into a convenient range. From the atomic point of view this softening of the glass by addition of soda results from the extra oxygen introduced by the soda, which produces single-bonded oxygens. Each single-bonded oxygen represents a break or weak spot in the strong three-dimensional silicon–oxygen network. The more of these breaks, the weaker the glass and the lower the temperature at which the glass can be worked.

VII. BORATE GLASSES

The oxide B_2O_3 readily forms a glass; in fact, crystallization can be produced only by taking very special precautions. The x-ray analysis (12) of vitreous boric oxide leads to the result that each boron is triangularly surrounded by three oxygens and each oxygen is bonded to two borons. Except for the lower coördination number, the same sort of random network is formed as in the case of vitreous silica.

In boric oxide–silica compositions (1) it is found that each boron is triangularly surrounded by three oxygens, each silicon is tetrahedrally surrounded by four oxygens, and each oxygen is bonded to two cations, either two borons, two silicons, or a boron and a silicon. The structure is a continuous three-dimensional random network. The addition of boric oxide to silica produces a softer glass by putting in boron, which bonds to the rest of the structure in only three directions, in place of silicon which bonds itself tetrahedrally in four directions to the surrounding atoms. A slight softening of silica glass by the addition of boric oxide is the basis of the familiar Pyrex chemical glasses.

In the soda–boric oxide (2) and the soda–boric oxide–silica glasses, a change in the coördination number of the boron atom is responsible for maxima and minima found in various physical properties. The boron atom can exist in either threefold triangular coördination or fourfold tetrahedral coördination, and is actually found in both kinds of bonding in crystalline borates. When soda is added to a glass containing boric oxide, the extra oxygen introduced by the soda furnishes the necessary extra oxygen for part of the borons to change to fourfold tetrahedral coördination. The effect of adding soda to a boric oxide glass is just opposite to the effect of adding soda to silica glass. When soda is added to silica, the coördination number of silicon remains 4, the extra oxygen produces single-bonded oxygens in the silicon–oxygen network, and the breaks in the network result in a weaker or softer glass. On the other hand, when soda is added to a boric oxide glass, over certain ranges of composition, the extra oxygen is taken up by the increase in the coördination number of the boron, and the increase in the number of bonds results in a strengthening or hardening of the glass.

VIII. DISCUSSION OF THE GLASSY STATE

The remarkable glass-forming properties of silica are due to two features which result from the coördination scheme. Even in the melt, there is probably a strong tendency for each silicon to surround itself by four oxygens, and hence for each oxygen to bond between two silicons. Although any such bonding must be continually forming and breaking, it will nevertheless put enough linkages into the melt to stiffen it up and to give the high viscosity which prevents the atoms from rearranging themselves in the orderly fashion necessary for crystallization. It is this factor which prevents crystallization on cooling. The fact that each oxygen is bonded to only two silicons puts such a flexibility into the scheme of structure that the random network is almost as stable as a crystalline arrangement. It is this flexibility which allows a random type of linkage to form and exist in the melt.

Zachariasen (13) has discussed the conditions for glass-forming ability among the oxides, and has pointed out the importance of the condition that each anion shall be bonded to not more than two of the glass-forming cations. The compounds SiO_2, GeO_2, and BeF_2 are ready glass formers, and all have the same structure in which the anion is bonded to two cations. Similarly in vitreous B_2O_3 the oxygens are bonded to only two cations. On the other hand, BeO is not a glass-forming oxide, because although the beryllium atom surrounds itself tetrahedrally by four oxygens, the composition requires that each oxygen be bonded to four beryllium atoms. Four bonds to each oxygen do not allow enough flexibility for a random network to be comparable in stability with the crystalline form. Enough flexibility in structure to allow the existence of a fairly stable random network is the essential requirement for glass-forming properties.

The oxides which are found in a commercial glass can be fairly well classified in two groups,—network formers and modifiers. The oxides SiO_2, B_2O_3, P_2O_5, and GeO_2 are the standard network formers. Oxides such as Na_2O, K_2O, CaO, BaO, and PbO modify the properties of the glass by changing the ratio of total oxygen atoms to network-forming cations.

In the organic field, glass-forming properties are most pronounced with materials such as sugar or glycerol, where the molecule contains a large number of hydroxyl groups. A random type of hydroxyl bonding between hydroxyl groups on neighboring molecules results in the necessary random linkages to produce a viscous melt. Here again it is the fact that there is a wide flexibility in the way in which the bonding can take place which allows the random type of network necessary for glass-forming properties.

REFERENCES

(1) Biscoe, J., Robinson, C. S., and Warren, B. E.: J. Am. Ceram. Soc. **22,** 180 (1939).
(2) Biscoe, J., and Warren, B. E.: J. Am. Ceram. Soc. **21,** 287 (1938).
(3) Bragg, W. H., and Bragg, W. L.: *The Crystalline State,* p. 189. Bell and Sons, London (1933).
(4) Debye, P.: Ann. Physik **46,** 809 (1915).
(5) Debye, P., and Menke, H.: Ergeb. techn. Röntgenkunde II, p. 1. Akademische Verlagsgesellschaft m. b. H., Leipzig (1931).
(6) Randall, J. T.: *Diffraction of X-rays and Electrons by Amorphous Solids, Liquids, and Gases.* John Wiley and Sons, Inc., New York (1934).
(7) Trostel, L. J.: J. Am. Ceram. Soc. **19,** 271 (1936).
(8) Valenkov, N., and Poray-Koshitz, E.: Z. Krist. **95,** 195 (1937).
(9) Warren, B. E.: J. Applied Phys. **8,** 645 (1937).
(10) Warren, B. E., and Biscoe, J.: J. Am. Ceram. Soc. **21,** 49 (1938).
(11) Warren, B. E., and Biscoe, J.: J. Am. Ceram. Soc. **21,** 259 (1938).
(12) Warren, B. E., Krutter, H. M., and Morningstar, O.: J. Am. Ceram. Soc. **19,** 202 (1936).
(13) Zachariasen, W. H.: J. Am. Chem. Soc. **54,** 3841 (1932).
(14) Zernike, F., and Prins, J. A.: Z. Physik **41,** 184 (1927).

THE BASIC CALCIUM PHOSPHATES AND RELATED SYSTEMS. SOME THEORETICAL AND PRACTICAL ASPECTS

SIDNEY EISENBERGER, ALEXANDER LEHRMAN, AND WILLIAM D. TURNER

Department of Chemical Engineering, Columbia University, New York, New York, and Department of Chemistry, The City College of the College of the City of New York, New York, New York

Received December 8, 1938

CONTENTS

Contributions to our knowledge of tricalcium phosphate, tetracalcium phosphate, and the apatites (for convenience, all three of these are here included in the term "basic calcium phosphates") have come principally from those concerned with the nature and deposition of the inorganic constituents of bone or in the utilization of phosphatic materials for plant food. However, the many exceedingly ingenious and difficult investigations of the basic phosphates to be found in the literature should be of interest to all who employ physicochemical methods and techniques, especially in connection with solid-phase phenomena. Some of the most

significant contributions have been made by mineralogists as a result of the occurrence in nature of many varieties of calcium phosphate. Moreover, the field of water purification and boiler feed water treatment is sadly in need of a fundamental understanding of the chemistry of these compounds, in view of the rising importance of processes involving the use or precipitation of calcium phosphate.

The considerable number of publications on the basic calcium phosphates attests first to the difficulties which are encountered and second to a failure to arrive at any substantial agreement on fundamentals. Since the application of the phase rule and x-rays to the study of these compounds, a great deal of useful and reliable data has been published. The purpose of this paper is to organize the available information and to present a conception of the nature of these compounds which offers a consistent explanation of their properties. It is hoped that the research necessary for clearing up some of the remaining questions in this field will be stimulated.

Part I. Calcium Phosphate Systems[1]

Studies of calcium phosphates encounter great practical difficulties. The principal sources of trouble in thermal investigations have been discussed by Trömel (218). They are (*a*) the reactivity of the compounds, (*b*) their sensitivity to reduction, and (*c*) the high temperatures required. These conditions eliminate utensils of ceramic materials, carbon, carbides, and most metals. Trömel reports that platinum has been successfully used by others up to 1600°C. and that he himself has found rhodium to be best suited for this work. Others have been able to use iron, nickel, or zirconium oxide crucibles, depending upon the temperature (24).

Investigations of aqueous systems, on the other hand, must contend with quite a different set of problems. The time required to establish equilibrium is often considerable (12, 36, 38, 39, 59, 82, 111, 117, 132, 191, 220, 224). Bassett, for example, reports that in some cases complete equilibrium is not attained at 25°C. in 19 months (12). The size of particles usually found in the solid phase is so minute as to create difficulties in settling and filtering precipitates, in obtaining clear x-ray powder photographs, and in the use of petrographic methods (11, 12, 25, 37, 54, 66, 74, 126, 220). The extremely low solubilities of the basic phosphates (11, 12, 39, 54, 72, 82, 83, 126, 153, 196, 197) make it difficult to obtain reliable analytical results for saturated solutions. The marked effect of the solutions on glass (11, 12) presents another problem. The magnitude of this effect is illustrated by an experiment performed by one of us in which the accumulation of silica in the solid phase after a week of refluxing in a Pyrex flask was 18 per cent of the mass of the solid. In

addition to the foregoing, one would suspect that the absorption of carbon dioxide by alkaline solutions, especially in the presence of excess calcium ions, would have to be guarded against. However, no mention is made of this factor.

Similar troubles have been observed for the basic calcium arsenates (44, 163, 164, 210) and for the phosphates of other metals (115, 116, 185).

I. THE BINARY SYSTEM $CaO-P_2O_5$

Much of the work done on this binary system has been beclouded by the failure to recognize the presence of a ternary system through the inclusion of small quantities of water (75). The difficulty with which

[1] The following tabulation of terminology and numerical values is appended for the convenience of the reader.

NAME	FORMULA	MOLE RATIO $CaO:P_2O_5$	WEIGHT RATIO $P_2O_5:CaO$	REMARKS
Dicalcium phosphate....	$CaHPO_4$	2:1	1.27 :1	
Pentacalcium phosphate.		5:2	1.01 :1	Existence doubtful
α-Tricalcium phosphate*...............	$Ca_3(PO_4)_2$	3:1	0.844:1	High-temperature form (called gamma variety in first paper by Bredig *et al.*)
β-Tricalcium phosphate*.........	$Ca_3(PO_4)_2$	3:1	0.844:1	Low-temperature form; transition temperature 1180°C.
Hydroxyapatite or hydroxylapatite.........	$Ca_{10}(OH)_2$-$(PO_4)_6$ or $3Ca_3(PO_4)_2 \cdot Ca(OH)_2$	10:3	0.760:1	Existence as a unique stoichiometric compound doubtful
Apatite.................	$3Ca_3(PO_4)_2 \cdot CaX_2$	10:3	0.760:1	X = CO_3, SO_4, F, Cl, OH, etc.
Tetracalcium phosphate*...............	$Ca_4P_2O_9$	4:1	0.633:1	

* The stable existence of either form of tricalcium phosphate or of tetracalcium phosphate in the presence of water is doubtful.

last traces of water are removed from the basic phosphates is illustrated by the work of Schleede, Schmidt, and Kindt (191), who ignited hydroxyapatite preparations and found that, in the absence of excess lime, the elimination of water was not complete under 1500°C. Trömel (218) has also shown that hydroxyapatite can be formed from binary mixtures by reaction with water vapor at temperatures as high as 1050°C! This stability of hydroxyapatite to high temperatures explains why Bassett and others always found their analyses falling short of 100 per cent even after the usual ignitions at about 900°C. This behavior was early dis-

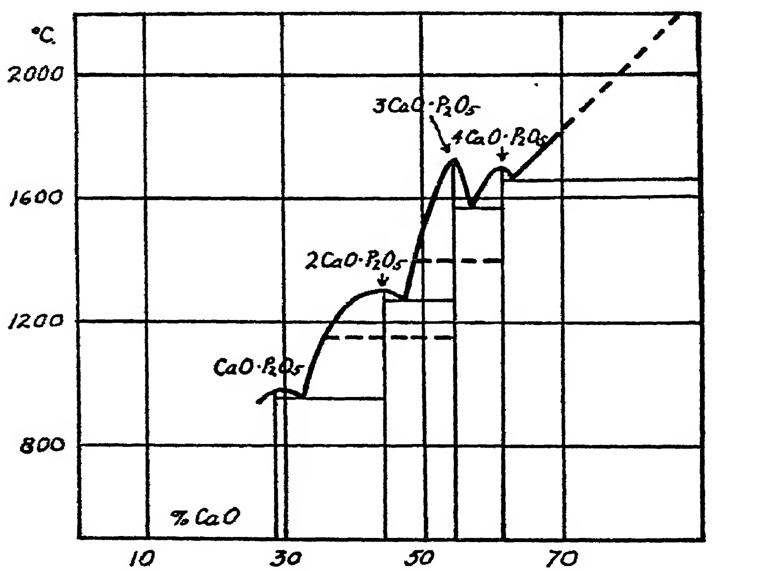


FIG. 1

covered, but the profound phase changes caused by traces of moisture have been only recently recognized.

The binary diagram worked out by Trömel is reproduced in figure 1. The temperature measurements were not precise enough (±20° at 1700°C.) to enable Trömel to be certain whether or not metaphosphate or pyrophosphate had congruent melting points, but that tri- and tetracalcium phosphates melted congruently was clearly indicated. The presence in the binary system of a molecular species with the formula $Ca_3(PO_4)_2$ has been thoroughly established by Trömel and his coworkers (119, 120, 217, 218, 219, 220) and by Bredig, Franck, and Füldner (24, 25). Schneiderhöhn (218) reports that $Ca_3(PO_4)_2$ exists in two enantiotropic forms with a transition point between 1250° and 1500°C. This is

confirmed in a paper published almost simultaneously by Bredig *et al.* (24), who locate the reversible transformation temperature at about 1180°C.[2]

Bredig *et al.* (25) have also shown that the alpha–beta transformation of tricalcium phosphate is greatly affected by the presence of moisture and excess calcium oxide. When the alpha-form, containing excess calcium oxide, was ignited in the absence of moisture, the beta lattice did not appear until the temperature had been lowered below 840°C., but in the presence of ordinary atmospheric water vapor, the conversion of alpha to beta occurred at the usual temperature of about 1200°C. The amount of water vapor required was remarkably small,—about 0.1

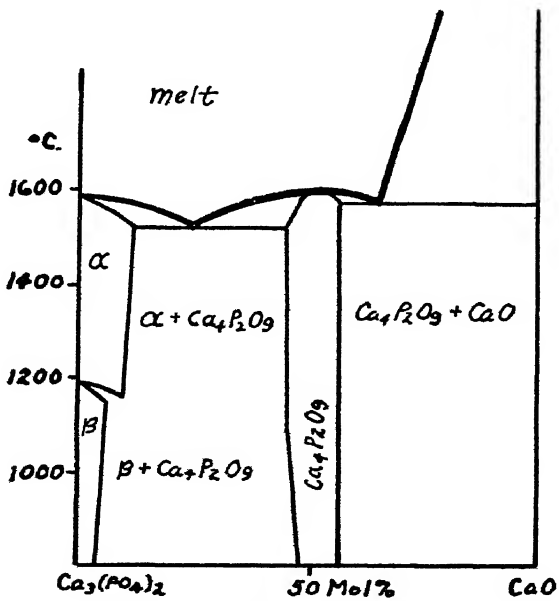


FIG. 2

to 0.2 per cent. α-Tricalcium phosphate showed no reluctance to transformation until excess calcium oxide was present. The authors concluded, therefore, that the alpha lattice was stabilized by calcium oxide which,

[2] In the binary system of Bredig *et al.* (24) the term "alpha phase" is applied to all compositions showing an apatite lattice, and the tricalcium phosphate stable at the higher temperature is called the gamma variety. In a later publication Bredig and his associates (25) agree with Trömel that the apatites have no place in the binary system and change their notation to conform with his. The variety of tricalcium phosphate which is stable at higher temperatures is called "alpha-tricalcium phosphate" by Trömel and his terminology will be followed in this paper.

however, could be eliminated by reaction with water and tricalcium phosphate to form hydroxyapatite.

The final diagram of Bredig *et al.* (25) is reproduced in figure 2 and shows almost perfect agreement with Trömel's diagram, except that some partial solubility in the solid phases is suggested. While the first diagram they proposed does not represent true equilibrium conditions, it does indicate the transformations which may be expected when ignitions are carried

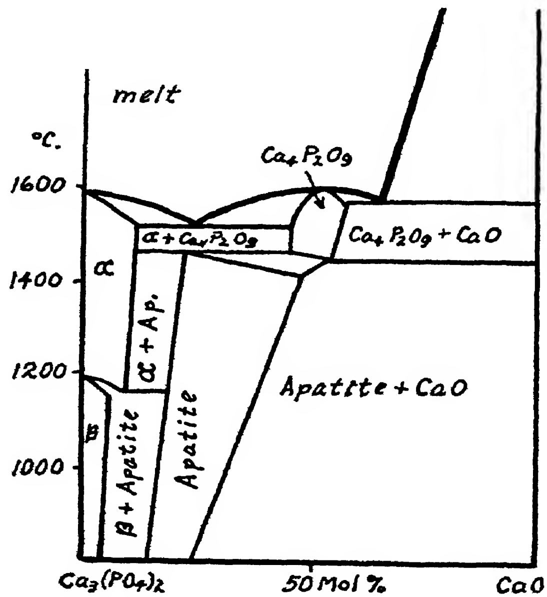


FIG. 3

out in the presence of moisture and hence is reproduced in figure 3. The following equations represent typical reactions:

$$3Ca_4P_2O_9 + H_2O \rightarrow Ca_{10}(OH)_2(PO_4)_6 + 2CaO\ ^3$$

$$Ca_{10}(OH)_2(PO_4)_6 \rightarrow 2Ca_3(PO_4)_2 + Ca_4P_2O_9 + H_2O\ ^4$$

$$3Ca_3(PO_4)_2 + CaO + H_2O \rightarrow Ca_{10}(OH)_2(PO_4)_6$$

[3] It is interesting to note that many years ago Foerster (57) discovered that 3 moles of $Ca_4P_2O_9$ gave on ignition 2 moles of CaO, which could be extracted with sugar solutions.

[4] Schleede *et al.* (191) have shown by means of x-ray powder photographs that when hydroxyapatite is heated to a temperature sufficient to remove the last traces of moisture a mixture of $Ca_3(PO_4)_2$ and $Ca_4P_2O_9$ is obtained.

The existence of oxyapatite, $Ca_{10}O(PO_4)_6$, has often been assumed (75). Since this compound has the composition $10CaO \cdot 3P_2O_5$, it should be found between $Ca_3(PO_4)_2$ and $Ca_4P_2O_9$, but all that can be discovered in this region is a well-defined eutectic of these compounds. Although this eliminates the possibility that oxyapatite can be crystallized from a melt, there still remains the possibility that it can be formed by reaction in the solid state. Trömel (218) explored this possibility by igniting mixtures corresponding to the composition of oxyapatite at various temperatures. However, no apatites could be found when water vapor was rigorously excluded from the ignitions. In addition to this phase rule demonstration of the non-existence of oxyapatite, McConnell (144) has shown that crystallographic considerations lead to the same conclusion.

It may be said in conclusion that the foregoing description of the binary system has a high degree of reliability. The data on which the conclusions are based were gathered through the correlated use of microscopic examination, chemical analysis, x-ray methods, and heating and cooling curves.[5] These four methods of approach gave mutually consistent results.[6]

II. THE BASIC REGION OF THE SYSTEM $CaO–P_2O_5–H_2O$

How can the nature of a precipitated material be determined? Where there is some assurance that the precipitate is a single solid phase or a mixture of two phases of known composition, a chemical analysis is sufficient. Unfortunately, in a great deal of work on the basic calcium phosphates, chemical analyses were relied on to determine the nature of precipitates without realizing that more than one solid phase could be present or without knowing what these solid phases were.

Where amorphous precipitates occur, as in the system under consideration, the difficulty of determining how many solid phases are present or of identifying the solid phases is obvious. The inadequacy of chemical analyses alone and the usefulness of the phase rule, especially in connection with the identification of amorphous basic salts, was early emphasized by Miller and Kenrick (149). They pointed out that if, in a ternary system, the compositions of precipitates are plotted against the compositions of the solutions in equilibrium with the precipitates, a horizontal line would

[5] Trömel reports (218) that calcium-rich mixtures (more than 50 per cent calcium oxide) gave no glasses even on chilling. The alpha- and beta-forms of $Ca_2P_2O_7$ and $Ca_3(PO_4)_2$ underwent transformation so slowly that heating and cooling curves were of no use in determining the transition temperatures. Since these points were only estimated by Trömel, the value determined for $Ca_3(PO_4)_2$ by Bredig, Franck, and Füldner (24) through x-ray examination of samples annealed at various temperatures is likely to be more accurate.

[6] Franck (60) and Trömel (221) have published reviews of much of the work described above.

indicate a monovariant system with a single, definite compound as the solid phase, a diagonal line would indicate a monovariant system with a solid solution of variable composition as the solid phase, and a vertical line would indicate an invariant system with two solid phases in varying proportions. This suggestion has become a standard procedure and there are innumerable examples of its application. However, there is no record of the use of this method for the basic calcium phosphates.

A large number of compounds more basic than dicalcium phosphate have been proposed solely on the basis of chemical analyses. The following is a partial list of such compounds: $Ca_4P_2O_9$ (11), $Ca_3(PO_4)_2$ (11, 12), $2CaHPO_4 \cdot Ca_3P_2O_8$ (163, 171), $5CaO \cdot 2P_2O_5 \cdot 10H_2O$ (55, 100, 186), and others (31, 171, 172). Some of these have been shown not to exist (74), and the others merit little consideration except insofar as supporting evidence is brought forward.

However, even when the phase rule is applied, certain elementary considerations must be taken into account. It is probably unnecessary to point out that the phase rule applies only to equilibrium conditions. But what does need emphasis is the danger of using time invariance of composition as a criterion for the attainment of equilibrium, especially where reaction rates are known to be very slow as they are for the basic calcium phosphates.

Another point which is sometimes neglected is the necessity for having homogeneous phases separated by sharp boundaries if the formula $P + F = C + 2$ is to be applicable. In dealing with amorphous solids, this condition is sometimes difficult to obtain, owing to surface effects such as adsorption.

The only certain criterion for equilibrium is to obtain identical results by approaching equilibrium from opposite sides. Moreover, if identical results are so obtained, it is virtually certain that adsorption effects are negligible. For if adsorption had taken place to any considerable extent, it is highly unlikely that its effect would be the same in both cases, considering the different environments on opposite sides of the equilibrium and the possibilities for particle size variation.

On the other hand, when adsorption does occur, the only certain method for detecting it and following changes in the solid phase is by means of x-rays, provided that changes in lattice dimensions are large enough to be detected.

There has been no fundamental investigation of the basic calcium phosphates which has given proper consideration to all of the foregoing factors. This undoubtedly accounts for the fact that the comparatively large amount of work done in this field is repetitious, disorganized, and contradictory.

Nevertheless, it is possible to bring some order into a seemingly bewildering situation and to resolve many of the present contradictions, if a few simple assumptions are made. *First, between dicalcium phosphate and lime, there exists, in the ternary system, a continuous series of solid solutions having an apatite lattice. It follows from this that tricalcium phosphate and hydroxyapatite do not exist in aqueous systems as unique, stoichiometric compounds.*[7] *Finally, in the words of McConnell (145), "The structure of apatite seems to be remarkably stable, permitting a number of unusual types of substitution and involving a considerable number of ions."*

These conclusions result from the critical review of the literature presented in the following sections.

A. The phase diagram

In 1908, Bassett (11) published a thorough phase rule study of the entire system at various temperatures. While his data for the isotherm at 25°C. were in good agreement with the results of previous workers (38, 39), his conclusions on the basic region were quite different. Whereas Cameron and his associates had given up the idea that tricalcium phosphate could be produced in an aqueous system in stable or metastable[8] form and had decided instead that dicalcium phosphate and lime formed a continuous series of solid solutions,[9] Bassett believed that the data showed the existence of tri- and tetra-calcium phosphates. He based his conclusion

[7] The terms "tricalcium phosphate" and "hydroxyapatite" are very widely and very loosely used. For example, some authors use the former for any precipitate more basic than dicalcium phosphate, although such precipitates have been frequently shown to have an apatite lattice or to be mixtures of dicalcium phosphate and an apatite. Others confine the use of the term to those precipitates with P_2O_5:CaO ratios approaching that of $Ca_3(PO_4)_2$. "Hydroxyapatite" is often used with the implication that the composition $Ca_{10}(OH)_2(PO_4)_6$, or several variations, has a unique existence as a definite compound. It would be tedious and repetitious to discuss in every case what term should have been used, and on the other hand confusing to make corrections without such discussion. Consequently, original terminology is used wherever the context is sufficiently clear to avoid confusion. The term "hydroxyapatite" will be used by us to mean any apatite belonging to the system $CaO–P_2O_5–H_2O$.

[8] Cameron (38) had originally suggested the possibility of the initial precipitation of metastable tricalcium phosphate with a slow rate of change to the stable form, but gave up the idea (39) when he could find no supporting evidence. Fouretier (59) also believed that the initial precipitate was metastable tricalcium phosphate which, however, reacted to form stable hydroxyapatite. The evidence for this belief is questionable.

[9] It is interesting to note that twenty-four years later Cameron (44) made a similar assertion for the basic region of the system $CaO–As_2O_5–H_2O$.

solely on the fact that analyses for some of his solid compositions were nearly correct for these compounds.

Since the data on which his conclusions for the basic region were based were rather meager, Bassett undertook a more extended investigation (12). Realizing the danger of relying on chemical analyses alone, Bassett plotted the concentrations of CaO against the concentrations of P_2O_5 in saturated solutions and looked for discontinuities in the curve to identify the solid phases. However, no discontinuities were found for tri- and tetra-calcium phosphates. While Bassett changed his mind about $Ca_4P_2O_9$[10], he reiterated his belief in the probable existence of tricalcium phosphate over a limited range of concentrations because so many of his preparations had compositions approaching the composition of $Ca_3(PO_4)_2$. He concluded from the failure to obtain a break in his curve at the boundary between di- and tri-calcium phosphates that there was no great difference in solubility between metastable dicalcium phosphate and stable tricalcium phosphate. He believed, moreover, that these small differences in solubility also accounted for the slowness of conversion of the former into the latter.

Bassett's data have been recalculated so that they could be plotted according to variations of the suggestion of Miller and Kenrick. In figures 4 and 5, the curves *ABC*, *AGH*, and *MEF* are consistent with Bassett's stated conclusions, but the curves *CD* and *HM*, which fit the data in a fairly satisfactory fashion, show that in the regions covered by them the solid phase is a solution of variable composition, having, as has been demonstrated by others, an apatite lattice (110, 220). Although some of the points are badly spotted, there seems to be no justification for a horizontal line at P_2O_5:CaO = 0.845, which should appear if tricalcium phosphate has a stable existence. From *D* and *M* to *E* the solid phase is undoubtedly hydroxyapatite. The solid phase along *EF* is reported by Bassett to be heterogeneous under the microscope and is evidently a mixture of varying quantities of hydroxyapatite and calcium hydroxide.

The stable existence of tricalcium phosphate in aqueous systems is highly improbable. Many others have come to a similar conclusion (36, 37, 38, 39, 54, 110, 112, 186, 187, 223).

The experimental results obtained by Jolibois (100) at 15–18°C. with a rapid mixing method (99) have been recalculated so that they could also be plotted in figures 4 and 5. The curve so obtained confirms the non-existence of tricalcium phosphate and the existence of a series of solid

[10] It has been pointed out (220) that the production of hydroxyapatite from $Ca_4P_2O_9$ by reaction with water in the atmosphere at 1050–1100°C. makes the existence of the latter in aqueous systems improbable.

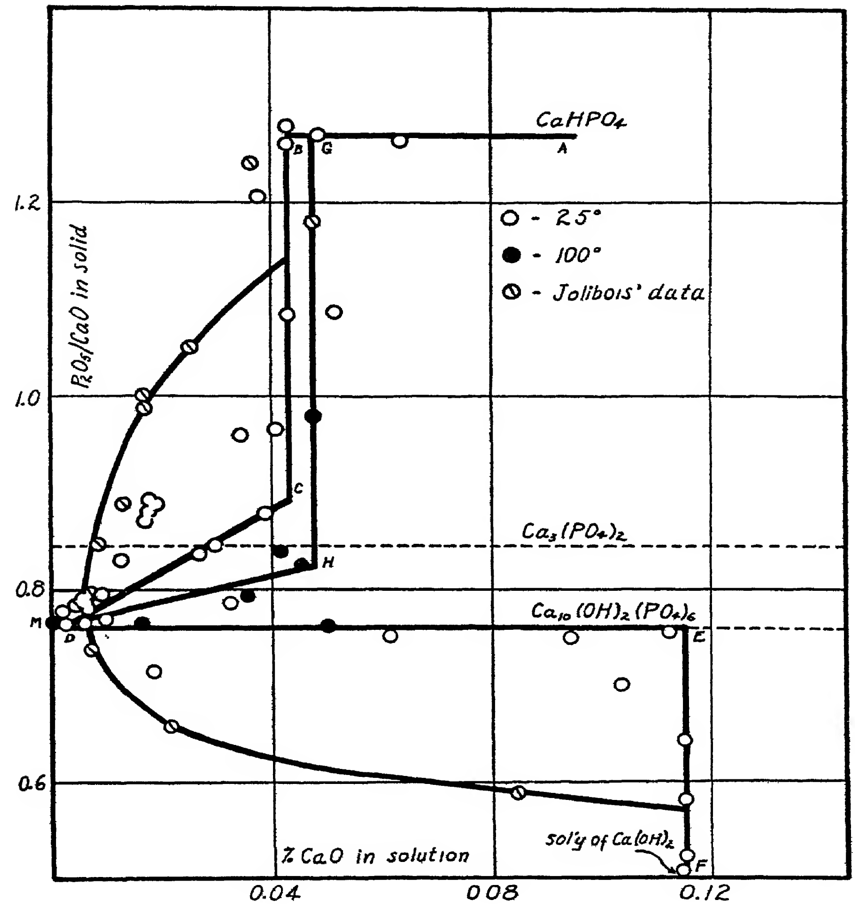


FIG. 4

solutions.[11] But Jolibois' results are different from Bassett's in one important respect: while the series of solid solutions disclosed by Bassett's data ranges in composition from dicalcium phosphate to hydroxyapatite,

[11] It may also be noted that there is no phase rule foundation for Jolibois' claim to the discovery of a new calcium phosphate. The mole ratios of CaO to P_2O_5 in the four compositions which he assumed to be pentacalcium phosphate varied considerably, while the solutions in contact with these solids had almost constant ratios. Such results are indicative of an invariant system with two solid phases and not of a monovariant system with a single solid phase. Jolibois reports that well-defined crystals may be observed, but furnishes no data on their physical and optical properties. One is forced to conclude that the crystals were probably dicalcium phosphate or pseudomorphs of dicalcium phosphate (12).

that of Jolibois extends to a ratio of P_2O_5 to CaO approximately equal to 0.58, with a solubility minimum at the composition of hydroxyapatite (P_2O_5:CaO = 0.76).

In either case, there is only one composition which is not decomposed by water.[12] The only solution which contains P_2O_5 and CaO in the same

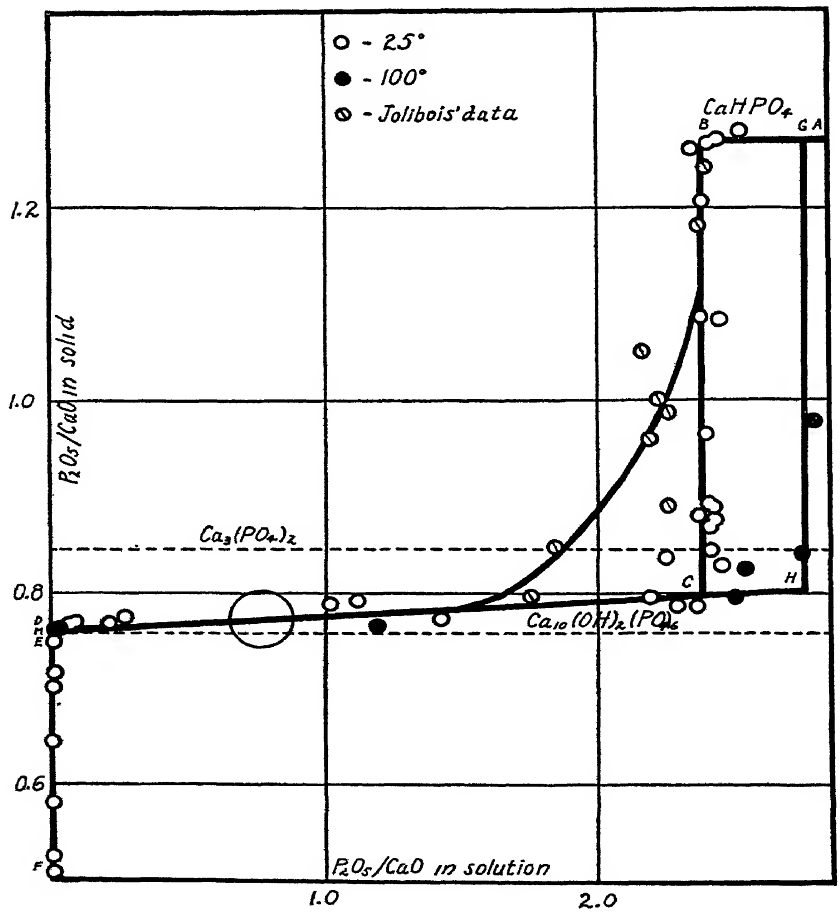


FIG. 5

proportion as is found in the solid phase in equilibrium with it is located somewhere in the circle in figure 5. Solids with larger ratios of P_2O_5 to CaO will become richer in CaO when treated with water (12, 31, 36, 37,

[12] Bassett's method of graphing results enabled him to demonstrate that mono-, di-, and tri-calcium phosphates were all incongruently soluble and that hydroxyapatite was the only calcium phosphate which was not decomposed by water. This

38, 40, 53, 57, 171, 183, 184, 186, 223),[13] while those with smaller ratios will become richer in P_2O_5 under similar treatment (18, 162). That hydrolysis proceeds in this manner has been completely verified by Schleede *et al.* (191). They found that all of the calcium phosphates will eventually display x-ray diffraction patterns similar to hydroxyapatite on protracted treatment with water. In addition, $Ca_4P_2O_9$ was the only one which gave alkaline solutions. In the face of this evidence, the following statements appearing in the literature must be incorrect: tricalcium phosphate hydrolyzes to $Ca_4P_2O_9$ (55); $Ca_4P_2O_9$ hydrolyzes to dicalcium phosphate and calcium hydroxide, and all phosphatic materials end as dicalcium phosphate on hydrolysis (227).

The chief uncertainty thus seems to be whether or not hydroxyapatite has a range of existence as a definite compound. The conception of an extended range of solid solutions suggested by Jolibois' data (but not by Jolibois himself) seems to furnish a more consistent explanation for the information available in the literature than does the assumption of a definite molecular species with the formula $Ca_{10}(OH)_2(PO_4)_6$.

For example, calcium phosphates with CaO contents in excess of the amount calculated for $Ca_{10}(OH)_2(PO_4)_6$ have been obtained by the following methods:

(*1*) Addition of limewater to dilute solutions of orthophosphoric acid. When Blarez (18) added excess quantities of limewater, he found that all of the precipitates had a mole ratio of CaO to P_2O_5 in excess of 3 and in some cases in excess of 3.3. Furthermore, all of the precipitates tended toward a composition of 3.6 moles of CaO per mole of P_2O_5 when allowed to stand in contact with their mother liquors whether they contained more or less than this amount immediately after precipitation.[14]

(*2*) Addition of solutions of calcium chloride to dilute alkaline solutions of sodium and potassium phosphates. Blarez reports results similar to those obtained when limewater and phosphoric acid were added. All of the precipitates approached a CaO content of 3.3 moles per mole of P_2O_5 when thoroughly washed with water.

(*3*) Extremely rapid mixing of limewater and orthophosphoric acid. The rapid mixing method of Jolibois produced precipitates (101) which approximated tricalcium phosphate in composition immediately after precipitation. However, on standing in contact with the mother liquor

---

confirmed the early results of Warington (223), who showed that persistent hydrolysis of precipitates of tricalcium phosphate caused them to approach the composition of hydroxyapatite.

[13] When calcium phosphates are repeatedly treated with neutral ammonium citrate solutions, the composition of the residues will also approach a ratio of P_2O_5 to CaO equal to about 0.76 (89).

[14] Berthelot (17) obtained similar results.

for 48 hr., the mole ratios of CaO to P_2O_5 increased to values between 3.48 and 4.50, depending on the amount of excess limewater used. Fouretier (59) found by means of x-rays that the initial precipitates were always a mixture of dicalcium phosphate and an amorphous solid. The final precipitates, however, always gave apatite patterns.

(*4*) Boiling precipitated tricalcium phosphate with concentrated sodium hydroxide. Foerster (57) obtained residues with 3.3 moles of CaO per mole of P_2O_5 when he used dilute sodium hydroxide, but with concentrated sodium hydroxide the ratio approached a value of 4.

(*5*) Boiling hydroxyapatite, prepared by ignition, with concentrated sodium hydroxide. Foerster also prepared various samples of hydroxyapatite by igniting mixtures of di- or tri-calcium phosphate with appropriate amounts of calcium oxide or calcium carbonate. In accordance with his previous results, these preparations were unaffected by dilute sodium hydroxide but approached a composition of 4 moles of CaO to 1 mole of P_2O_5 in concentrated caustic solutions. Foerster was able to obtain even 5 to 1 ratios by treating $Ca_4P_2O_9$ or Thomas slags with concentrated sodium hydroxide, but these residues probably contained some free lime.

(*6*) Treatment of hydroxyapatite, prepared by hydrolysis with water or dilute sodium hydroxide, with solutions of limewater. Lorah, Tartar, and Wood (132) found that the lime content of the residues increased steadily for a long time and that the amount taken up at any given time depended on the concentration of lime in solution.

In view of the differences in particle size undoubtedly found in these experiments and differences in solution environment, adsorption cannot furnish a complete explanation of the similarity of results (101, 132). It is equally improbable that large amounts of calcium oxide in the solid phase could be due to free lime, since methods 1, 3, and 6 would then indicate that the solubility of lime is materially lowered in the presence of hydroxyapatite. That this is not so may be seen in figure 4, where it is apparent that mixtures of hydroxyapatite and free lime are in equilibrium with solutions having the same concentration of calcium oxide as saturated limewater. Only one conclusion remains:—solid solutions with more CaO than $Ca_{10}(OH)_2(PO_4)_6$ have a stable existence in contact with solutions of varying concentrations (102). Blarez's results (18) are particularly striking. While he did not achieve equilibrium, it is obvious that he approached a condition of equilibrium from opposite directions.

B. The phase diagram for calcium arsenates

The phase diagram for calcium arsenates has been the subject of some recent studies. The evidence is briefly presented because of the similarities between the phosphates and the arsenates.

The results obtained by Pearce and Norton (163) for the 90°C. isotherm are plotted in figure 6 on an enlarged scale. It will be noted that penta- and tri-calcium arsenates, as well as the analog of hydroxyapatite, are assumed to exist. A similar type of curve is reported for 35°C. (164), but no evidence is found for the basic compound and the ranges of equilibrium concentrations in solution are somewhat different.

The method used was ingenious. It consisted of arranging the concentrations of CaO and As_2O_5 in the solution so that when the temperature was raised to the desired level only a small amount of solid would precipitate (calcium arsenates are less soluble at higher temperatures). If the

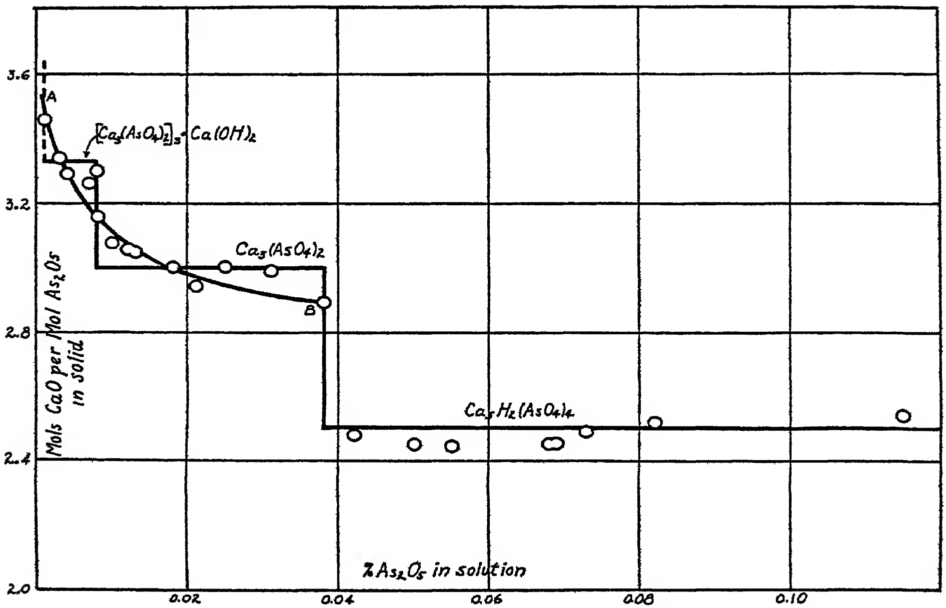


FIG. 6

separation of the solid could take place with little disturbance to conditions in solution, the initial precipitate would be so near to its final composition that equilibrium would be reached in a very short time.

However, if the ratios of CaO to As_2O_5 in the solutions before and after precipitation are calculated from the data of Pearce and Norton, it is found that while the change in this ratio is generally of the order of 10 per cent, several of the most basic compositions show a much larger change,—up to 700 per cent.

It may be assumed, therefore, that the method used did not yield equilibrium conditions in the most basic region. That the curve reported by Pearce and Norton for this region is probably incorrect, is shown by

the fact that it was necessary for them to assume that one of the preparations consisted of a mixture of the basic compound and free lime. It would be surprising if precipitation from limewater produced an equilibrium with free lime in the solid phase. The curve *AB* in figure 6 has been drawn to represent a series of solid solutions. It may be seen that it fits the data at least as well as the curve drawn by the original authors, and the point which originally was taken to indicate a mixture of basic arsenate and lime in the solid phase now has a more reasonable significance.

A more recent publication (156), in which x-ray and petrographic methods were used to determine the 62°C. isotherm, confirms the existence of a solid solution in the basic region. Some effort was made to determine the extent of solid solution on the high-lime side but it appears, unfortunately, that the limiting composition on the low-lime side was assumed to be tricalcium arsenate. The existence of pentacalcium arsenate, reported by Pearce and his associates, has also been confirmed by Nelson and Haring and seems to rest on much sounder evidence than does the existence of pentacalcium phosphate. The hydrolysis of the calcium arsenates follows much the same course as does that of the calcium phosphates.

C. The nature of precipitated basic calcium phosphate

Attention has been concentrated thus far mainly on evidence susceptible to phase rule treatment. However, because of the practical difficulties inherent in phase rule procedures, many investigators have turned to x-ray diffraction methods. But despite the fundamental character of such research, a number of different conclusions have been proposed.

Trõmel and Möller (220) found that a number of precipitates of different composition gave x-ray patterns identical with that of hydroxyapatite.[15] Since they could detect no change in lattice dimensions as the composition varied, they concluded that precipitated tricalcium phosphate was essentially hydroxyapatite with sufficient adsorbed phosphate to yield the proper composition and not an hydrato-apatite, $Ca_9(H_2O)_2(PO_4)_6$, as proposed by Hendricks *et al.* (74).[16]

[15] Precipitation in boiling solutions improved the sharpness of the lines in the x-ray photographs. Trõmel and Möller took this as an indication of growth in particle size. An alternative explanation,—one which many x-ray workers ignore and which is very applicable to this field,—is a growth in uniformity of composition among the various particles. If particles differing in composition but all having an apatite lattice are present, the small differences in lattice dimensions will cause the lines to lose sharpness.

[16] Hendricks *et al.* believed that this hydrato-apatite could be distinguished from other apatites by the disappearance of the apatite pattern on ignition. From the discussion in the section on the binary system, it is evident that this is not strictly

Trömel and Möller made much of the fact that the composition of the precipitate when the calcium solution was poured into the phosphate solution was different from the composition when the order of mixing was reversed (66, 127). They believed that in the former case an opportunity for adsorption is offered and accepted. However, it has been amply demonstrated that initial precipitates of basic calcium phosphates require much time to change over to their equilibrium compositions (page 258). Moreover, it has already been pointed out that adsorption cannot furnish a complete explanation for such phenomena (page 270).

The failure to obtain any noticeable differences in the x-ray patterns of widely varying compositions is probably due to the smallness of the change in lattice constants and insufficient precision in x-ray methods. That this is plausible is shown by the fact that Bredig *et al.* (25) have prepared a calcium aluminate- and a calcium ferrite-apatite which cannot be distinguished from fluoro-apatite by x-rays. It is to be expected, however, that with recent improvements in precision (45, 46, 47, 222) such slight differences will become measurable.

Bredig *et al.* (25) offered an entirely different concept to explain the continued appearance of apatite patterns as the composition was varied,—the formation of mixed apatites (apparently what was meant was something like double or multiple salts). They were able to show that $Ca_4P_2O_9$ could be transformed into an apatite at high temperatures by the absorption of water in quantities much less than the stoichiometric amount for hydroxyapatite. They approached the same result from the opposite direction by removing part of the water from hydroxyapatite. The product was considered to be a mixed hydroxy-oxy apatite. In a similar manner, a mixed fluoro-oxy apatite was synthesized by absorption of small quantities of calcium fluoride under conditions eliminating the effect of moisture. The existence of other so-called mixed apatites was postulated, and Bredig *et al.* were very successful in explaining certain anomalous phenomena with them.

All of the foregoing authors recognized inadequacies in their hypotheses. Hendricks *et al.* (74) concluded, "For structural reasons such a compound (hydrato-apatite) would be expected to form a complete series of solid solutions with hydroxyapatite . . ." Trömel and Möller (220) suggested the possibility of the formation of mixed crystals (presumably synonymous with solid solutions) of hydroxyapatite and hydrated tricalcium phosphate.

so, since any hydroxyapatite will lose the apatite pattern if ignited in the absence of moisture. The same may also be true of the so-called carbonate–apatites. Others have shown that precipitated tricalcium phosphate can have an apatite pattern or the pattern of either of the anhydrous tricalcium phosphates, depending on the temperature of drying (24, 220).

Bredig *et al.* (25) thought it probable that precipitated tricalcium phosphate was composed of mixed crystals of hydrated tricalcium phosphate (in which the water does not necessarily follow any stoichiometric relationship) and hydroxyapatite (with or without adsorbed phosphate according to circumstances), but their final conclusion was that it is not yet possible to come to any definite decision. The essential difficulty seems to be a certain reluctance to abandon the idea of discrete molecular types among the apatites.

The results of this unwillingness to accept the idea of a crystal lattice existing without discrete molecular types are also illustrated by the work of Roseberry, Hastings, and Morse (178). They came to the conclusion that tricalcium phosphate may exist as a definite, independent crystal form. And yet they felt it necessary to suggest that this crystal form appears to be the nucleus of the apatite series. When CaX_2 is associated with tricalcium phosphate, its molecules are placed in the crystal lattice where they cause no detectable difference in the important planes. They believed that tricalcium phosphate is not as stable a form as $CaX_2 \cdot nCa_3(PO_4)_2$, where X = Cl, F, $\frac{1}{2}O$, or $\frac{1}{2}CO_3$ and n is not less than 2 nor more than 3.

Larson (126) has more recently proposed a monohydrate formula, $Ca_3(PO_4)_2 \cdot H_2O$, for precipitated tricalcium phosphate. The evidence is not convincing. The x-ray patterns for the freshly precipitated substance and its product of ignition are reported to be different from the oxy- and hydroxy-apatite patterns of Trömel (217). This is undoubtedly true for the ignited salt, which should give the pattern for β-tricalcium phosphate. The failure to obtain or to recognize an apatite pattern for the un-ignited material is probably due to the generally unsatisfactory nature of the basic precipitates for x-ray powder work when freshly precipitated. The remaining evidence is the loss of 0.97 mole of water on ignition at 950–970°C.

D. Reaction rates

Figure 7 is reproduced from a paper by Buch (31), who extended Rindell's work on the hydrolysis of dicalcium phosphate. Rindell (172) states that equilibrium is not attained in some cases in 252 hr. Schleede *et al.* (191) also report that $CaHPO_4 \cdot 2H_2O$, agitated for 80 hr. with constantly renewed hot water, gave x-ray patterns of $CaHPO_4$ and did not show the apatite pattern until after 400 hr. of hydrolysis. It is obvious, therefore, that the 20-hr. extractions on which figure 7 is based did not give equilibrium conditions and do not prove, as Buch thought, that compounds exist with compositions between those of dicalcium and tricalcium phosphates.

Buch's results are significant, nevertheless, in that they lead to some

idea of reaction rates. From figure 7 it would seem that the rate of hydrolysis of a mixture of dicalcium phosphate and solid solution decreases until a critical proportion of the latter is present. Sanfourche and Henry (184) found that a small quantity of "tricalcium phosphate" was necessary to induce dicalcium phosphate to hydrolyze. With the critical quantity of solid solution present, the rate rises rapidly and remains fairly constant until the dibasic salt is no longer present. Buch states that his residues did not become completely amorphous (indicating complete disappearance of dicalcium phosphate) until just before the sharp drop from the upper plateau in figure 7. The rate of reaction from then on is extremely slow and decreases as the residue becomes richer in calcium oxide.

It is instructive to determine the results which Buch and Rindell should have obtained had they been able to reach equilibrium in their

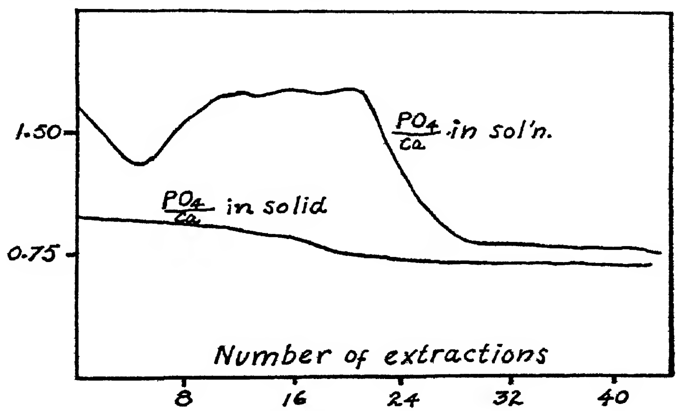


FIG. 7

extractions. In figure 8 point *A* represents pure dicalcium phosphate. On treatment with water, the solution will contain relatively more phosphate than the original solid. Hence at equilibrium, represented by point *B*, the solid will have proportionately more calcium oxide. Separation of the solid from the solution brings the system to point *C*. A second extraction ends at point *D*, which represents a solid phase in equilibrium with a solution of the same composition as the one which was in equilibrium with solid B. Thus, as long as dicalcium phosphate persists in the solid phase, the composition of the solution after each extraction will remain constant. This is illustrated by *BD* in figure 9. However, as soon as the system becomes monovariant, each successive solution will display continuously decreasing ratios of P_2O_5 to CaO, until finally at *E* the proportion becomes equal to that found in the solid. At this point

the solid ceases to become richer in calcium oxide and all succeeding solutions will have identical compositions.

III. THE STABILITY OF THE APATITE LATTICE

The apatite lattice is tolerant not only to large variations in the components of the ternary system $CaO–P_2O_5–H_2O$, but also to the inclusion

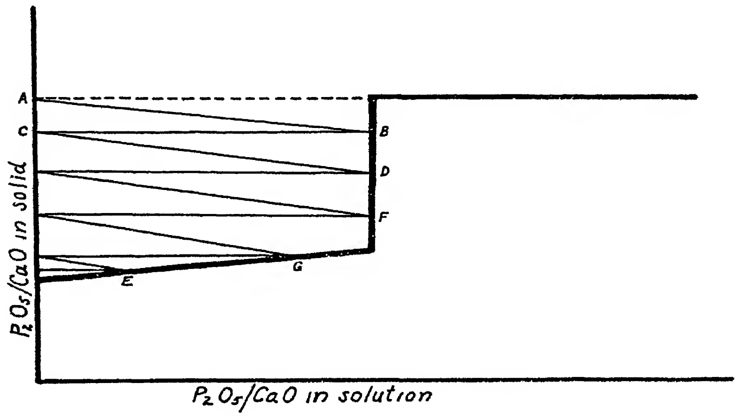


FIG. 8

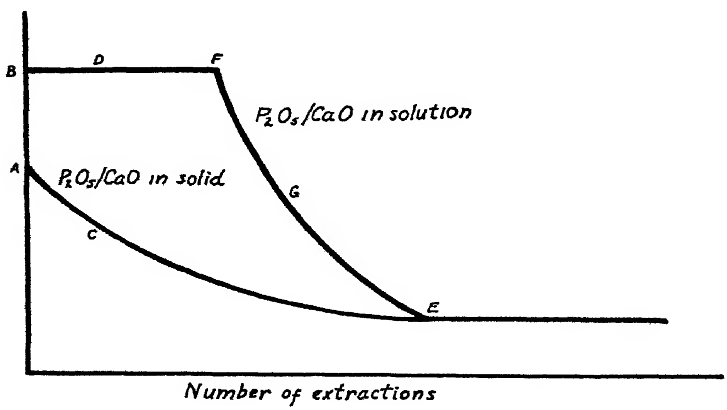


FIG. 9

of many other substances in large or small amounts. From the evidence which is about to be examined, one is inevitably forced to the conclusion, already stated for the ternary system, that *it is only a fortuitous occurrence when the composition of any apatite may be expressed by small whole number ratios of atomic species suggesting a definite chemical compound.*

A. Mineralogical studies

Schleede *et al.* (191) have stated, "According to the lattice determinations of Schiebold and Mehmel (147) and of Náray-Szabó (181), relatively large spaces are available for fluorine atoms or hydroxyl groups . . . . As a result . . . still larger units such as Cl (155), CO_3 (23, 56), or SO_4 (23, 56) may be built in. It does not appear excluded, then, that a still larger number of molecules of water besides the two hydroxy groups can be placed in the space provided."[17, 18]

The most revealing investigations of apatite minerals are those of McConnell and his associates (68, 69, 144, 145, 146). On the basis of x-ray studies of carefully prepared and carefully analyzed[19] samples, the following conclusion was reached (69): "Each unit cell (of fluorapatite) possesses the following types of ionic positions: 24 O, a few of which may be occupied by OH or F ions, if an excess of either should be present; all of the O ions are bonded to P ions; 2 F, which may be entirely, or in part, occupied by OH, Cl, or O ions; 6 P, a small number of these positions,

[17] Our translation.

[18] The structures proposed by Mehmel and by St. Náray-Szabó have been reproduced several times (68, 69, 74, 145). Lattice dimensions are given in the following table:

	a_0	c_0	c/a	VOLUME	DENSITY BASED ON MOLECULAR WEIGHT OF 1008.9
	Å.	Å.		cu. Å.	
Fluoro-apatite (145)	9.36 ±0.01	6.88 ±0.01	0.7350	522	3.187
Fluoro-apatite (181)	9.37 ±0.01	6.88 ±0.01	0.7342	523	3.180
Fluoro-apatite (147)	9.36 ±0.02	6.85 ±0.02	0.7318	520	3.201
Francolite (68)	9.34 ±0.01	6.88 ±0.01	0.7366		
Dahllite (69)	9.41	6.88	0.7311	528	
Enamel (69)	9.41	6.87	0.7301	527	3.055
Dentine (69)	9.40	6.87	0.7309	526	3.024
Hydroxyapatite (34)	9.42	6.94	0.736	533	
Hydroxyapatite (116)	9.40	6.93	0.737		
Chloro-apatite (116)	9.52	6.85(?)	0.719		
Lead hydroxyapatite (116)	9.90	7.29	0.736		
Pyromorphite (116), $Pb_{10}Cl_2(PO_4)_6$	9.95	7.32	0.736		

Optical properties of apatite minerals have been published by others (3, 52).

[19] McConnell has pointed out that much of the previous work is unreliable for crystallographic analyses because of the difficulty in securing accurate chemical analyses. This is substantiated by the scheme worked out by Harvey (71) for the complete analysis of apatite rock.

probably not exceeding 10 per cent, may be occupied by C, V, or As; Si and S may also occupy these positions; 10 Ca, a small amount of Ca may be replaced by C, in which case all of the ten positions are not necessarily filled, because C can replace more than one Ca, depending upon the number of charges required to produce electrostatic equilibrium." In addition to these isomorphous substitutions, the following were also suggested (145): for calcium,—magnesium, manganese, strontium, sodium, potassium, and possibly barium, chromium, iron, aluminum, cerium, and other rare earths.[20] Where oxygen substitutes for fluorine, either (*a*) part of the calcium is replaced by a trivalent metal, or carbon, or possibly nitrogen; or (*b*) part of the phosphorus is replaced by a hexavalent ion such as sulfur.

Other chemical, optical, and x-ray studies of phosphatic minerals have established the fact that the apatite lattice is capable of enduring very wide variations in composition (21, 74, 75, 174, 175, 176, 189, 190, 207).[21]

B. Basic phosphates of other metals

Klement (116) has recently discovered some remarkable similarities between the basic regions of the ternary systems CaO–P_2O_5–H_2O and PbO–P_2O_5–H_2O. For example, the final product of the hydrolysis of dilead phosphate is the analog of hydroxyapatite. It is asserted, moreover, that lead hydroxyapatite is the only basic lead phosphate stable in water.

Klement (115) has also made a thorough investigation of the basic phosphates of magnesium. Although dimagnesium phosphate does not hydrolyze in the same way as dicalcium phosphate, compositions such as $Mg_{10}(OH)_2(PO_4)_6 \cdot 27H_2O$ can be obtained by hydrolysis in pure water. Since this substance is monoclinic instead of hexagonal, it is not an apatite. Klement's data suggests the possibility that a continuous series of solid solutions also exists in this system.

Sanfourche (185) believes that strontium forms an analog of hydroxyapatite but that barium does not. This conclusion is based on certain similarities in behavior when the alkaline-earth hydroxides are titrated with phosphoric acid and should be regarded as tentative until additional information is obtained.

Jolibois and Cloutier (102), using the former's rapid mixing method, have concluded that the basic phosphates of lead, copper, iron, aluminum, zinc, silver, manganese, chromium, and uranium probably do not exist. This conclusion is also in need of confirmatory evidence, especially since it has already been shown to be wrong with respect to lead.

[20] Körber and Trömel (119) have prepared an apatite compound with the formula $La_2Ca_8(PO_4)_6 \cdot O_2$.

[21] For detailed optical properties and a bibliography up to 1929, see Hausen (73).

Part II. Practical Applications

Enough is now known of the aqueous calcium phosphate system so that one can begin to understand materials and processes of great practical importance. However, a word of caution is necessary at the very beginning. It has not been easy to learn what we know of the system CaO–P_2O_5–H_2O, and it is by no means certain that our knowledge is complete or correct in all of its details. But the soil, mineral, and biological systems which are the subjects under consideration in this section are infinitely more complex in their relationships, and therefore our knowledge of them is much less reliable. It is unfortunate that many workers in these fields have given insufficient consideration to the difficulties of applying the simple things that can be done in the laboratory to complicated natural phenomena. The conclusions, which have been reached on a foundation in which many factors, some of controlling importance, have been ignored, are rash and have tended to distract attention from much good data. It is obvious, therefore, that much of what is stated here is of the nature of a first approximation and more dependable knowledge waits upon future work.

Natural processes are dynamic, and information based on studies of equilibrium conditions is necessarily of limited usefulness. It is doubtful that we shall get very far with investigations of soil reactions or bone deposition merely with what is known of ionic equilibria, solubility product, or simple phase rule systems. Future work should emphasize investigations into phenomena in flux rather than systems in their final states. Only then will our present knowledge achieve its maximum usefulness.

I. PHOSPHATIC FERTILIZERS

Since this review is confined to the basic calcium phosphates, the manufacture and composition of superphosphate or triple superphosphate will not be presented here. However, attention should be drawn to the excellent paper of Hill and Hendricks (77), in which a combination of chemical, optical, and x-ray methods has been used to determine the composition of these materials. For economic information, see Jacob's reports (93).

A. Calcined phosphate

The availability of $Ca_4P_2O_9$ for plant food has long been known, but it was not until comparatively recently that α-$Ca_3(PO_4)_2$ was also shown to be highly available (25, 91, 215). It follows then from the discussion of the binary system CaO–P_2O_5, that if fluorine could be removed from phosphate rock, the phosphate content could be made available by a simple ignition at 1200°C. or higher, followed by rapid cooling, or by cooling in a dry atmosphere. Providing that the presence of silica

has no marked effect, such a procedure should produce a mixture of α-$Ca_3(PO_4)_2$ and $Ca_4P_2O_9$ from most phosphate rocks.

The major problems are (*a*) how to remove fluorine, as the presence of fluorine has a marked effect in rendering phosphatic materials unavailable (4, 5, 10, 25, 104, 120, 134, 180); (*b*) how to prevent or minimize the formation of hydroxyapatite through reaction with atmospheric moisture during cooling; (*c*) how to prevent or minimize the transformation of α-tricalcium phosphate to the less soluble β-tricalcium phosphate during cooling. The Bureau of Chemistry and Soils of the United States Department of Agriculture made a systematic effort to solve these problems, and the results are contained in a series of papers which are remarkable for their thoroughness and soundness (95, 142, 166, 167, 168).

The problem of removing fluorine was solved by high-temperature ignitions in the presence of (*a*) steam and silica, or (*b*) water vapor alone, or (*c*) silica alone with the effectiveness varying in the order given. There was no significant increase in citrate-solubility at any temperature until all the fluorine had been removed except the amount theoretically required for the formula $Ca_{10}(OH)(F)(PO_4)_6$. In fact, the removal of less than 65 per cent of the fluorine generally caused a decrease in citrate-solubility, confirming an earlier observation of Bredig *et al.* (25) that fluoro-oxy-apatite had a smaller citrate-solubility than fluoro-apatite. After the removal of 65 per cent of the fluorine, the citrate-solubility depended jointly on the amount of fluorine further volatilized and the temperature. If the volatilization was carried out at 1350°C. or higher, or if the temperature is subsequently raised to that level, the increase in citrate-solubility was roughly proportional to the amount of fluorine further removed. Once the fluorine was volatilized,—and this could be done at comparatively lower temperatures,—the ignition temperature was the controlling factor, in that before the charge was chilled a temperature of at least 1350°C. had to be attained.

It was found that if the furnace was cooled slowly to 1300°C. and then the charges chilled by withdrawal from the furnace, no significant decrease in availability occurred. However, if slow cooling was continued to temperatures lower than 1200°C., the decrease in availability was large. If in the latter case the annealing was performed in a dry atmosphere, the decrease in citrate-solubility was materially lessened although it was still considerable.

By means of a carefully planned series of experiments, it was shown that the presence of silica is beneficial for the conversion of most varieties of synthetic and natural phosphates to forms available for plant food. This is not unexpected, since Körber and Trömel (119, 120) have shown that, in the system CaO–P_2O_5–SiO_2, α-$Ca_3(PO_4)_2$ and $Ca_4P_2O_9$ have ex-

tended ranges of homogeneity and the velocity and temperature of the transformation of α-tricalcium phosphate to β-tricalcium phosphate are considerably reduced. Furthermore, if silicophosphates are formed, their P_2O_5 contents are highly available.

Marshall *et al.* (142) have suggested the following mechanism for the calcining process: (*a*) one atom of fluorine from fluoro-apatite is replaced by hydroxyl, yielding hydroxyfluoro-apatite which is very insoluble in neutral ammonium citrate; (*b*) the second atom of fluorine is replaced by hydroxyl, yielding hydroxyapatite which also has a low citrate-solubility; (*c*) hydroxyapatite decomposes. Water vapor and silica are necessary for the first two steps. Water interferes with the third step unless silica is present.[22]

Subsequent x-ray investigations (78, 143) disclose that the rôle played by silica is not a simple one and that reversion, or decrease in availability, involves several complex phase transformations.

While the desirable effects of dry annealing over wet annealing were, in general, confirmed, marked variations were found and seem to be due to variable composition. The transformation of α-$Ca_3(PO_4)_2$ to the beta variety does not take place and hence cannot account for reversion. However, reversion was always accompanied by the appearance of apatite and is almost certainly due solely to the formation of apatite. The evidence seems to indicate that two different apatites are produced, one which requires the presence of moisture and one which does not. The latter forms at a higher temperature and usually has a greater effect on citrate-insolubility. Contrary to expectations, the presence of $Ca_4P_2O_9$ was not reported. While α-$Ca_3(PO_4)_2$ is usually the dominant phase in the unreverted samples, silicocarnotite and one or more phases of unknown composition may also occur.

The advantages of calcined phosphate over superphosphate have been summed up (142) as follows: (*a*) the product is a sintered or semi-fused clinker requiring no aging, merely grinding; (*b*) the product is practically insoluble in water and gives weakly alkaline solutions, thereby avoiding bad effects on machinery and fertilizer bags, although possibly leading to loss of ammonia when mixed with ammonium salts (13); (*c*) the product contains about 30 per cent available P_2O_5 as compared with 19 to 20 per cent in the best grades of superphosphate (transportation cost is a considerable item in the total cost to the consumer (48)); (*d*) the product has superior mechanical condition alone and in mixtures (226); (*e*) pot tests (30, 94, 179, 198) show plant food values as high as superphosphate or dicalcium phosphate; (*f*) low fluorine content creates the possibility of

[22] Messerschmitt (148) believes that the silica combines with any excess calcium, thus preventing the formation of apatite.

substitution for bone meal in mineral feeds for livestock. The process cannot be used for phosphate rocks rich in iron or aluminum phosphate, since these are rendered almost completely unavailable by calcination.

An effort at commercial development (48) has, however, uncovered the following practical difficulties: (*a*) the temperature required is just below the fusion temperature and incipient fusion leads to problems of handling; (*b*) phosphate rocks vary considerably in composition, producing varying fusion temperatures, and therefore close temperature control becomes necessary. A good product was obtained by fusing the rock and bubbling in dry steam. The objections to this procedure were low refractory life[23] and high fuel costs. These obstacles seem to have been overcome by the St. Jacques turbulent furnace (182), in which the fine particles of rock are suspended in a gas stream during the reaction period. A commercial plant based on the use of this furnace has been constructed in Algeria.

Until recently, the calcining process seemed destined to play an important part in the phosphate fertilizer industry (93, 229). It may still do so in other countries, but in the United States there has been a sharp drop in interest. This is probably due to the remarkable results reported by the Tennessee Valley Authority on calcium metaphosphate (50, 135), which contains better than 60 per cent available P_2O_5. Here may be an example of a very promising process becoming obsolete even before it could be placed into large-scale production. It would be particularly striking to find such an example in this industry, in which technological progress has lagged so far behind.[24]

B. Thomas meal

If the discussion on calcined phosphates is broadened to include mixtures richer in lime and silica, then the basic slags of the steel industry become included. Tetracalcium phosphate and one or more silicophosphates are the important sources of available P_2O_5 in these slags (4, 120, 194).

The compositions of the silicophosphates are still uncertain. The following formulas have been proposed: (*a*) $5CaO \cdot P_2O_5 \cdot SiO_2$ (4, 14, 19, 24, 119, 120, 124, 191, 193) (silicocarnotite); (*b*) $9CaO \cdot P_2O_5 \cdot 3SiO_2$ (119, 120); (*c*) $8CaO \cdot P_2O_5 \cdot 2.5SiO_2$ (157); (*d*) $7CaO \cdot P_2O_5 \cdot 2SiO_2$ (157); (*e*) $10CaO \cdot 3P_2O_5 \cdot SiO_2$ (57); (*f*) $3CaO \cdot P_2O_5 \cdot 3SiO_2$ (161); (*g*) $13CaO \cdot 3P_2O_5 \cdot SiO_2$ (124). Formulas *e* and *g* are not supported by adequate experimental data and *f* has been ruled out by x-ray examinations (24). Compositions *b*, *c*, and *d* have identical diffraction patterns and are thus one and the same com-

[23] For results on commercial operation of phosphate-smelting furnaces, see reports by the Tennessee Valley Authority (No. 51,160).

[24] Problems connected with analyses of calcined phosphate have been worked out by Jacob and his associates (96, 179).

pound. It may be said then that only two ternary compounds have been discovered in the system $CaO–P_2O_5–SiO_2$. And only on the formula of silicocarnotite has there been any general agreement. The readiness with which solid solutions are formed has much to do with the difficulty in determining the compositions of the various compounds in this system (119).

Much of what has been said of calcined phosphate may be applied to Thomas meal. As far back as 1892, Foerster (57) observed that a large portion of the P_2O_5 in $Ca_4P_2O_9$ became unavailable on annealing at moderately high temperatures. He very shrewdly ascribed the effect partly to temperature transformation and partly to reaction with constituents in the atmosphere. On these assumptions he recommended that basic slags be cooled rapidly,—a bit of advice which has been repeated more recently (104, 191).

Fluorine in basic slags is just as objectionable as fluorine in calcined phosphates. Thus, efforts to increase the amount of fertilizer by-product by adding raw phosphate rock to molten slags have failed because of the effect of the fluorine in the added material on the availability of the P_2O_5 (120).

II. REMOVAL OF FLUORIDE ION FROM WATER

The stability of the apatite lattice to a large variety of isomorphous replacements creates the possibility of developing a number of useful ion-exchange reactions.

The recognition of fluorine as the cause of the disease known as "mottled teeth" has stimulated the development of processes for removing small quantities of fluorides from potable waters. Several recent proposals are based on the use of calcium phosphates. The usefulness of such compounds is undoubtedly due to the ability of fluoride ion to replace hydroxyl ion in the apatite lattice.

The removal of solute fluoride by the formation of fluoro-apatite was apparently first noted by MacIntire and his associates (133, 134). The first actual attempt to develop a practical process seems to have been made by Smith and Smith (208), who used ground bone as the source of replaceable hydroxyl ion.

MacIntire and Hammond (136) have shown that almost any method whereby hydroxyapatite is produced in the water to be purified[25] is effective. For economic reasons, however, it is better to have the water flow through towers packed with the active substance and to devise a method for regenerating exhausted material.

[25] For example, baking powder suspensions made alkaline with ammonia or calcium hydroxide.

The replacement of OH^- by F^- is easily reversed by washing with dilute sodium hydroxide, but the subsequent removal of excess alkali was not so easy. A water wash followed by a wash with dilute hydrochloric acid was found to be effective (2, 208) but was open to the following objections (15): metal corrosion; loss of phosphate; and an effective life of only about twenty-five cycles. These difficulties were overcome by substituting a water solution of carbon dioxide for the hydrochloric acid (15).

Patent protection for these processes has been sought (2, 15, 230).

III. THE INORGANIC CONSTITUENTS OF BONE[26]

Although the idea of isomorphous replacement in the apatite lattice has already been tentatively proposed in connection with the nature of the inorganic constituents of bone (6, 178, 211), it has made surprisingly little headway. Now that Gruner *et al.* (69) have published a detailed study of the apatite lattice in teeth, giving an exact description of the various possibilities of isomorphous substitution, there should be little excuse for retaining certain outworn conceptions. It is hoped that the weight of evidence here assembled will stimulate a general reorientation on the problem of bone deposition.

X-ray studies have established beyond any question the apatite nature of the salts of bones and teeth (6, 7, 8, 25, 26, 43, 69, 74, 103, 110, 111, 112, 117, 150, 151, 169, 170, 178, 188, 211, 212, 213, 214).

Marek, Wellman, and Urbányi (137, 138, 139, 140, 141) discounted x-ray evidence on the ground that conflicting conclusions had been reached. They failed to recognize, however, that whatever other disagreements x-ray investigators have had, they all agree that the crystallites in bones and teeth are apatites of one kind or another. The only exception is Funaoka (62), who noted the similarity between the x-ray patterns of bones and precipitated tricalcium phosphate but was not aware that the latter was an apatite.

Marek *et al.* also pointed out that the crystal structure of the samples used for x-ray analyses have been so altered by chemical or thermal treatment as to render them no longer representative of the original bone or tooth substance. It is true that many investigators have ignored the possibility that essential changes might occur when bones are ignited or subjected to the action of a glycerol solution of potassium hydroxide,—both common methods for removing organic matter. When Marek and his associates demonstrated (139) that the usual wet methods for isolating pure mineral constituents produce modifications in composition, they

[26] For information on bony structures as a whole and an excellent bibliography, consult Huggins' review (88). Reviews are available for other aspects of calcification (173, 192).

emphasized the dangers of such neglect. But, on the other hand, it is also true that much evidence has been produced for which this criticism is not valid. Apatite patterns have been reported for a large variety of animal and human bones, teeth, and pathological calcification where the samples have been untreated or have been freed of their organic matter by methods which do not appear to have any effect on the essential inorganic structure (26, 74, 169, 178, 211).

It should be clear, therefore, that any attempt to resolve the contradictory views on the nature of the mineral constituents of bone must be limited to apatite structures.

Since ". . . the crystalline constitution of enamel and dentine (and bones) is not entirely uniform and may vary considerably from person to person, and even from one part of the tooth to another, and still not be perceptible in appearance or even in such crystal diffraction photographs as have hitherto been made" (6), the chemical evidence must be examined to determine whether or not particular apatites are present in bones and teeth (111, 213).[27]

A. Carbonate–apatite versus hydroxyapatite

Werner and others (32, 33, 64, 205, 225) came to the conclusion that carbonate–apatite is the principal inorganic constituent of bone, because bone analyses often approximated the proper proportions of calcium, phosphate, and carbonate. Others have shown that the carbonate content of bone cannot be due to the presence of free calcium carbonate.

Gassmann (64) pointed out that calcium carbonate was easily soluble in glacial acetic acid, while the carbonate in bone was not.[28] He (64, 65) also heated bones and teeth with barium chloride and found that chlorine became bound in some insoluble form. He believed that if free calcium carbonate were present, soluble calcium chloride would have been formed.[29]

Hendricks and coworkers (74) heated mixtures of calcium carbonate and "hydrated tricalcium phosphate" or "pure" hydroxyapatite to con-

[27] Many years before x-rays were known or applied in this field, the composition of bony substances and their chemical behavior led to the assumption that an apatite was present (1, 85, 225). Numerous analyses show an approximately constant mole ratio of Ca:P = 10:6 for many kinds of bone under a wide variety of conditions (20, 22, 28, 29, 63, 64, 80, 81, 87, 107, 122, 137, 152, 195, 200, 204, 205, 209, 225).

[28] Klement (108) disputed this finding by showing that calcium carbonate is practically insoluble in anhydrous glacial acetic acid. It must be recognized, however, that whatever reagent Gassmann used, it had a different effect on calcium carbonate than it did on bone.

[29] Klement has remarked (108) that the chlorine could have been rendered insoluble by reaction with hydroxyapatite even if free calcium carbonate were present.

stant weight at 900°C.; in both cases free lime almost equivalent to the added calcium carbonate was formed. Similar treatment of bone samples gave no substantial amount of free lime.

Thus there seems to be little reason to doubt that carbonate in bone is part of the fundamental apatite.[30] The contention of Klement (107, 111, 117) that the absence of a 10:6:1 ratio of calcium, phosphate, and carbonate in saturated solutions of bone proves the absence of carbonate–apatite and that the presence of a comparatively large amount of carbonate in solution (131) proves the presence of hydroxyapatite and the carbonates or bicarbonates of sodium, potassium, and magnesium (12) is without merit. For if complex apatites exist in bone, it is very likely that they are incongruently soluble[31] and can establish equilibria with solutions of widely varying compositions.

However, a great deal of analytical data is available to show that the inorganic constituents of bone cannot be represented by the formula $Ca_{10}(CO_3)(PO_4)_6$.

It has been demonstrated that bone contains more basic equivalents than acid equivalents, leading to the supposition that hydroxyl groups make up the difference (63, 107, 117, 128, 152). Furthermore, the proportion of carbonate is quite variable and seems to depend on definite factors such as age, disease, and species (28, 58, 87, 105, 117, 121, 125, 128, 137, 138, 152, 158, 159, 202). In many cases, after deducting from the total calcium amounts equivalent to the carbonate present, the ratios of residual calcium to phosphorus are greater than the theoretical ratio for tricalcium phosphate and sometimes even exceed the theoretical ratio for hydroxyapatite (121, 152, 202).

Not only is the proportion of carbonate often less than would be expected for $Ca_{10}(CO_3)(PO_4)_6$, but it is occasionally more (121, 202).

Thus, those who consider carbonate–apatite the principal inorganic constituent of bone (64, 65) are no more right than those who support hydroxyapatite for that rôle (118).

B. Minor constituents

Evidence on the forms taken by the minor constituents of bone is comparatively meager and in general inconclusive. There are some reasons

[30] Henschen, Straumann, and Bucher (76) have reported lines for calcium carbonate in x-ray spectrograms of bone, but their results have never been duplicated by any of the other x-ray investigators.

[31] The term "incongruently soluble" is usually applied to a compound which is partially or completely changed to another compound in the process of forming a saturated solution. It is used here to signify any change in solid phase, especially the formation of a solid solution of different composition, caused by the fact that no equilibrium exists between a given solid and a solution of identical solute composition.

for believing, however, that magnesium, alkali metals, chlorine, and fluorine are part of the main apatite crystal.

As far back as 1872, Aeby (1) suggested that fluorine occurs as a substitute for oxygen in oxyapatite, since varying amounts had little, if any, effect on the relative proportions of calcium and phosphorus. The fluorine content is variable within a considerable range but is always much less than the theoretical for fluoro-apatite (22, 63, 97, 98, 105, 106, 110, 113, 114, 117, 199, 216). The minerals in fossil bone usually contain increased amounts of fluorine (42, 165, 175, 176, 177), which appear to enter the lattice by isomorphous replacement of CO_3 and H_2O (74) or OH (114).

Analytical results form the basis of several attempts to account for magnesium and the alkali metals as carbonates or bicarbonates (12, 107, 110, 111, 117). These attempts fail for many bone compositions (138, 139, 154). Moreover, Logan (128) makes the pertinent observations that alkali metals, which do not form insoluble phosphates or carbonates, are deposited in calcified structures from low concentrations in blood plasma and that the failure of glycerol solutions of potassium hydroxide (70) to extract the alkali metals is not comprehensible unless they are integral components of the inorganic structure. The comparative ease with which part of the carbon dioxide may be removed from bone by heat (12) does not prove the presence of free carbonates, since it has been shown that synthetic apatites containing carbonate lose perceptible amounts of carbon dioxide at comparatively low temperatures.

Klement has recently concluded that magnesium may replace calcium isomorphously in bone apatite (115).

C. Crystal orientation

According to a number of x-ray studies, the crystallites in enamel are preferentially orientated (6, 41, 43, 169, 188, 212, 213, 214), while those in dentine have a random arrangement (6, 41, 43, 212). Clark (43) reports that specimens of bone and enamel which he has examined have apatite crystals orientated in fibers. Thewlis (214) states that the orientation on the surface of enamel is different from the orientation in enamel sections.

There has been little effort, as yet, to determine the significance of crystallite orientation, except that Reynolds *et al.* (169) believe that one of the effects of fluorine in mottled teeth is to inhibit orientation and that this is one of the factors causing brittleness in fluoride teeth.

D. Physicochemical considerations

From time to time, efforts have been made to apply the methods of physical chemistry and the latest concepts of heterogeneous equilibria to the study of the mechanism of bone deposition. But these researches

have, in the main, contributed little of value, because some of the restrictions imposed on thermodynamic methods have been violated. For example, the concept of solubility product is meaningless without reference to a known solid phase of comparatively simple ionic composition. Yet it has been stated (154) that "The generally accepted view is that calcification takes place in calcifying tissues from solutions rendered saturated or supersaturated when the product ($Ca^{++} \times PO_4^{---}$) exceeds the solubility product. The material deposited, 'the bone salt,' is an insoluble phosphate of the apatite series $3Ca_3(PO_4)_2 \cdot CaX_2$." Perhaps this is an exaggerated statement, but it is indicative of a tendency to use a concept without proper regard for its limitations (129, 130, 196).

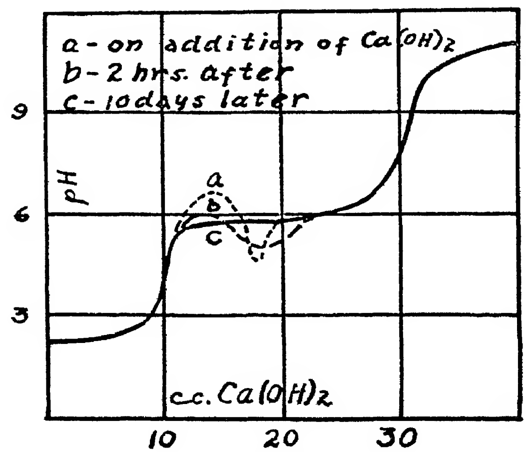


FIG. 10

Holt and his associates did attempt to identify the solid phase (82) before calculating solubility products (83, 84). They titrated phosphoric acid with limewater and confirmed the results of Wendt and Clarke (224), which are reproduced in figure 10.[32] They believed that the coincidence of the upper inflection point with the addition of exactly enough calcium hydroxide to form the tertiary compound indicated the existence of tricalcium phosphate. This compound, however, is almost certainly not capable of stable existence in aqueous solutions, especially at the pH of blood serum or the pH of the inflection point (12).

That smooth titration curves could be obtained in 8 to 10 days is further confirmation of the point made earlier that dicalcium phosphate reacts comparatively rapidly with solutions from which it precipitates.

[32] Britton (27) obtained similar results, but those of Hoffman and Gortner (79) were somewhat different.

But once it has been completely converted to a more basic form, the speed of reaction between solid and solution becomes very small. Thus, it is probable that the upper inflection point is due merely to rapid readjustment of ionic equilibria only slightly complicated by slow reactions with the solid phase.

Kramer, Shear, and associates have shown time and again, *in vivo* and *in vitro*, that the product Ca $\times$ P in solution is a criterion for calcification at the pH of serum (86, 121, 123, 201, 202, 203). Furthermore, in equilibration experiments with dicalcium phosphate at various pH values (202), the product, $[Ca^{++}] \times [HPO_4^{--}]$, calculated from theoretical considerations previously developed by others (72, 82, 83, 84, 196, 197), was found to be substantially constant. They suggested the interesting theory that dicalcium phosphate is the first substance precipitated in calcification and that the final basic compounds are obtained by subsequent hydrolysis (109, 201). Klement (111, 117) has recently supported this theory. However, there is no evidence of any kind that this compound occurs in bone. As a matter of fact, when Kramer *et al.* (121) investigated the composition of primary calcification (in which dicalcium phosphate, pure or partially hydrolyzed, should be found if it is to be found anywhere), they discovered that instead of the ratio of calcium to phosphorus being less than normal, it was considerably more.

Considering the complexity of the solid phase and the solution from which it precipitated, it seems hopeless to undertake research based on the assumption that "... initiation of precipitation of the bone salt depends alone on the concentrations of calcium, phosphate, and hydrogen ions" (130). The following statement by Murray (154) presents a more promising approach to the problem: "The fact that big differences in mineral content can occur in material deposited in different regions of the body at the same time promotes the suggestion that all calcifications are the result of specific cell activity rather than merely precipitation governed by solubility products."

Conclusion

According to Walden and Cohen (222), the criteria for the formation of solid solutions in ionic crystals which have been usually applied[33] "... may perhaps be valid for the occurrence of complete miscibility, but their force is certainly weak in the case of limited solubility." This observation seems to be supported by the apatite family, which furnishes an example of an ionic crystal capable of dissolving a large variety of

[33] Identical types of chemical structure, identical types of crystal structure, and nearly identical lattice parameters.

components, in large or small proportions, without fundamental change in lattice.

Furthermore, it is contended that, at least for hydroxyapatite, a substance with some small whole number ratio of atomic species does not exist but rather a crystal lattice common to a continuous series of solid solutions.

The present uncertainties in our knowledge call for a new phase rule investigation of the basic region of the system $CaO–P_2O_5–H_2O$. Lengthy periods of contact between solid and solution should not be the sole method for obtaining equilibrium conditions. The method of Pearce and his associates (163, 164) may be adapted to this purpose, or precipitations may be carried out in concentrated solutions of inert salts in the hope that increased solubility will speed up the rate of reaction. Ammonium acetate is suggested, since it may easily be eliminated before analysis of the liquid phase. Carrying out reactions at high temperatures in bombs is another possibility.

But whatever method is used, time invariance of composition ought not to be the sole criterion of equilibrium. Equilibrium should be approached from opposite directions. This is possible, since it has been shown that compositions may be prepared which are either richer in CaO or richer in P_2O_5 than equilibrium compositions merely by altering the mode of precipitation.

Now that increased precision in x-ray methods is available (45, 46, 47, 222), a number of possibilities suggest themselves. The range of existence of solid solutions in the ternary system or in more complex apatite systems may be ascertained. The rôle played by the various constituents of bone and teeth may be established by determining the effect of changing composition on the lattice dimensions of the fundamental mineral matter.

REFERENCES

(1) Aeby, C.: J. prakt. Chem. **5**, 308 (1872).
(2) Adler, H., Klein, G., and Lindsay, F. K.: Ind. Eng. Chem. **30**, 163 (1938).
(3) Alfani, M.: Mineralog. Abstracts **5**, 283.
(4) Bainbridge, F.: Iron and Steel Inst. (London), Carnegie Schol. Mem. **10**, 1 (1920).
(5) Bainbridge, F.: Trans. Faraday Soc. **16**, 302 (1921).
(6) Bale, W. F., Hodge, H. C., and Warren, S. L.: Am. J. Roentgenol. Radium Therapy **32**, 369 (1934).
(7) Bale, W. F., and Hodge, H. C.: Naturwissenschaften **24**, 141 (1936).
(8) Bale, W. F., LeFevre, M. L., and Hodge, H. C.: Naturwissenschaften **24**, 636 (1936).
(9) Bartholomew, R. P., and Jacob, K. D.: J. Assoc. Official Agr. Chem. **16**, 598 (1933).

(10) BARTHOLOMEW, R. P.: Arkansas Agr. Expt. Sta., Bull. 312, 20 (1934); Soil Sci. **40**, 203 (1935).
(11) BASSETT, H., JR.: Z. anorg. Chem. **59**, 1 (1908).
(12) BASSETT, H., JR.: J. Chem. Soc. **111**, 620 (1917).
(13) BEESON, K. C., AND JACOB, K. D.: Ind. Eng. Chem. **30**, 304 (1938).
(14) BEHRENDT, G., AND WENTRUP, H.: Arch. Eisenhüttenw. **7**, 95 (1933).
(15) BEHRMAN, A. S., AND GUSTAFSON, H.: Ind. Eng. Chem. **30**, 1011 (1938).
(16) VAN BEMMELEN, J. M.: Z. anorg. Chem. **15**, 90 (1897).
(17) BERTHELOT: Compt. rend. **132**, 1277 (1901).
(18) BLAREZ, C.: Compt. rend. **103**, 264 (1886).
(19) BLOME, H.: Métallurgie **7**, 659, 698 (1910).
(20) BOGERT, L. J., AND HASTINGS, A. B.: J. Biol. Chem. **94**, 473 (1931).
(21) BORNEMAN-STARYNKEVICH, I. D.: Compt. rend. acad. sci. U. R. S. S. **19**, 253 (1938).
(22) BOWES, J. H., AND MURRAY, M. M.: Biochem. J. **29**, 102, 2721 (1935).
(23) BRAUNS, R.: Neues Jahrb. Mineral. Geol. Beilage Bd. **41**, 60 (1917); J. Chem. Soc. **110**, II, 532 (1916).
(24) BREDIG, M. A., FRANCK, H. H., AND FÜLDNER, H.: Z. Elektrochem. **38**, 158 (1932).
(25) BREDIG, M. A., FRANCK, H. H., AND FÜLDNER, H.: Z. Elektrochem. **39**, 959 (1933).
(26) BREDIG, M. A.: Z. physiol. Chem. **216**, 239 (1933).
(27) BRITTON, H. T.: J. Chem. Soc. **1927**, 614.
(28) BROOKE, R. O., SMITH, A. H., AND SMITH, P. K.: J. Biol. Chem. **104**, 141 (1934).
(29) BROUWER, E.: Biochem. Z. **190**, 402 (1927).
(30) BROWN, B. E., REID, F. R., AND JACOB, K. D.: Am. Fertilizer **81**, No. 13, 5, 27 (1934).
(31) BUCH, K.: Z. anorg. Chem. **52**, 325 (1907).
(32) BURNS, C. M., AND HENDERSON, N.: Biochem. J. **29**, 2385 (1935).
(33) BURNS, C. M., AND HENDERSON, N.: Biochem. J. **30**, 1207 (1936).
(34) BURRI, C., JAKOB, J., PARKER, R. L., AND STRUNZ, H.: Schweiz. mineralog. petrog. Mitt. **15**, 327 (1935).
(35) BYLER, W. H.: J. Am. Chem. Soc. **60**, 1247 (1938).
(36) CAMERON, F. K., AND HURST, L. A.: J. Am. Chem. Soc. **26**, 885 (1904).
(37) CAMERON, F. K., AND SEIDELL, A.: J. Am. Chem. Soc. **26**, 1454 (1904).
(38) CAMERON, F. K., AND SEIDELL, A.: J. Am. Chem. Soc. **27**, 1503 (1905).
(39) CAMERON, F. K., AND BELL, J. M.: J. Am. Chem. Soc. **27**, 1512 (1905).
(40) CAMERON, F. K., AND BELL, J. M.: U. S. Dept. Agr. Bull., 41 (1907).
(41) CAPE, A. T., AND KITCHEN, P. C.: J. Am. Dent. Assoc. **17**, 193 (1930).
(42) CARNOT, A.: Ann. mines **3**, 155 (1893).
(43) CLARK, J. H.: Am. J. Physiol. **98**, 328 (1931).
(44) CLIFFORD, A. T., AND CAMERON, F. K.: Ind. Eng. Chem. **21**, 69 (1929).
(45) COHEN, M. U.: Rev. Sci. Instruments **6**, 68 (1935).
(46) COHEN, M. U.: Z. Krist. **94**, 288 (1936).
(47) COHEN, M. U.: Z. Krist. **94**, 306 (1936).
(48) CURTIS, H. A., MILLER, A. M., AND JUNKINS, J. N.: Chem. & Met. Eng. **43**, 583 (1936).
(49) CURTIS, H. A., COPSON, R. L., BROWN, E. H., AND POLE, G. R.: Ind. Eng. Chem. **29**, 766 (1937).

(50) Curtis, H. A., Copson, R. L., and Abrams, A. J.: Chem. & Met. Eng. **44**, 140 (1937).
(51) Curtis, H. A., Miller, A. M., and Newton, R. H.: Chem. & Met. Eng. **45**, 116 (1938).
(52) Dadson, A. S.: University of Toronto Studies, Geol. Series No. 35, 51 (1933).
(53) Danneel, H., and Fröhlich, K. W.: Landw. Jahrb. **72**, 51 (1930).
(54) Danneel, H., and Fröhlich, K. W.: Z. Elektrochem. **36**, 302 (1930).
(55) Danneel, H., and Fröhlich, K. W.: Z. anorg. allgem. Chem. (Schenk-Festschrift) **188**, 14 (1930).
(56) Eitel, W.: Schriften königsberg. gelehrten Ges., Naturw. Klasse, **1924**, 168.
(57) Foerster, O.: Z. angew. Chem. **1892**, 13.
(58) Forbes, E. B., Halverson, J. O., Morgan, L. E., Schulz, J. A., Wells, E. B., Hunt, C. H., and Winter, A. R.: Ohio Agr. Expt. Sta. Bull. 347, 99 pp. (1921).
(59) Fouretier, G.: Compt. rend. **205**, 413 (1937).
(60) Franck, H. H.: Congr. chim. ind., 14th Congr., Paris, Oct., **1934**, 13 pp.
(61) Franck, H. H., Bredig, M. A., and Frank, R.: Z. anorg. allgem. Chem. **230**, 1 (1936).
(62) Funaoka, S.: Ber. ges. Physiol. exptl. Pharmakol. **63**, 588 (1932).
(63) Gabriel, S.: Z. physiol. Chem. **18**, 257 (1894).
(64) Gassmann, T.: Z. physiol. Chem. **55**, 455 (1908); **63**, 397 (1909); **70**, 161 (1910); **83**, 403 (1913); **90**, 250 (1914); **178**, 62 (1928); **185**, 234 (1929); **192**, 61 (1930); **201**, 284 (1931).
(65) Gassmann, T.: Ber. **70B**, 41 (1937).
(66) Glass, N., and Jones, A. J.: Quart. J. Pharm. Pharmacol. **5**, 442 (1932).
(67) Greenwald, I.: J. Biol. Chem. **124**, 437 (1938).
(68) Gruner, J. W., and McConnell, D.: Z. Krist. **97**, 208 (1937).
(69) Gruner, J. W., McConnell, D., and Armstrong, W. D.: J. Biol. Chem. **121**, 771 (1937).
(70) Harrison, H. E.: J. Biol. Chem. **120**, 457 (1937).
(71) Harvey, C. O.: Analyst **61**, 817 (1936).
(72) Hastings, A. B., Murray, C. D., and Sendroy, J., Jr.: J. Biol. Chem. **71**, 723 (1927).
(73) Hausen, H.: Acta Acad. Aboensis, Math. Phys. **5**, No. 3, 62 pp. (1929).
(74) Hendricks, S. B., Hill, W. L., Jacob, K. D., and Jefferson, M. E.: Ind. Eng. Chem. **23**, 1413 (1931).
(75) Hendricks, S. B., Jefferson, M. E., and Mosley, V. M.: Z. Krist. **81**, 352 (1932).
(76) Henschen, C., Straumann, R., and Bucher, R.: Z. Chirurgie **236**, 485 (1932).
(77) Hill, W. L., and Hendricks, S. B.: Ind. Eng. Chem. **28**, 440 (1936).
(78) Hill, W. L., Hendricks, S. B., Jefferson, M. E., and Reynolds, D. S.: Ind. Eng. Chem. **29**, 1299 (1937).
(79) Hoffman, W. F., and Gortner, R. A.: J. Phys. Chem. **29**, 769 (1925).
(80) Holmes, A. D., and Pigott, M. G.: Ind. Eng. Chem. **23**, 190 (1931).
(81) Holmes, A. D., Pigott, M. G., and Campbell, P. A.: J. Biol. Chem. **92**, 187 (1931).
(82) Holt, L. E., Jr., LaMer, V. K., and Chown, H. B.: J. Biol. Chem. **64**, 509 (1925).
(83) Holt, L. E., Jr., LaMer, V. K., and Chown, H. B.: J. Biol. Chem. **64**, 567 (1925).

(84) HOLT, L. E., JR.: J. Biol. Chem. **64,** 579 (1925).
(85) HOPPE, F.: Arch. path. Anat. (Virchow's) **24,** 13 (1862).
(86) HOWLAND, J., AND KRAMER, B.: Trans. Am. Ped. Soc. **34,** 204 (1922).
(87) HOWLAND, J., MARRIOTT, W., AND KRAMER, B.: J. Biol. Chem. **68,** 721 (1926).
(88) HUGGINS, C.: Physiol. Rev. **17,** 119 (1937).
(89) JACOB, K. D., HILL, W. L., ROSS, W. H., AND RADER, L. F., JR.: Ind. Eng. Chem. **22,** 1385 (1930).
(90) JACOB, K. D., AND ROSS, W. H.: Com. Fertilizer Yearbook, **1932,** 19, 46.
(91) JACOB, K. D., RADER, L. F., JR., MARSHALL, H. L., AND BEESON, K. C.: Ind. Eng. Chem., Anal. Ed. **4,** 25 (1932).
(92) JACOB, K. D., RADER, L. F., JR., AND ROSS, W. H.: J. Assoc. Official Agr. Chem. **15,** 146 (1932).
(93) JACOB, K. D.: Mineral Ind. **43,** 448 (1934); **44,** 452 (1935); **45,** 471 (1936); **46,** 459 (1937).
(94) JACOB, K. D., BARTHOLOMEW, R. P., BROWN, B. E., PIERRE, W. H., REID, F. R., AND TIDMORE, J. W.: J. Agr. Research **50,** 837 (1935).
(95) JACOB, K. D., REYNOLDS, D. S., AND MARSHALL, H. L.: Am. Inst. Mining Met. Engrs., Tech. Pub. 695, 14 pp. (1936).
(96) JACOB, K. D., RADER, L. F., JR., AND TREMEARNE, T. H.: J. Assoc. Official Agr. Chem. **19,** 449 (1936).
(97) JODLBAUER, A.: Z. Biol. **41,** 487 (1901).
(98) JODLBAUER, A.: Z. Biol. **44,** 259 (1903).
(99) JOLIBOIS, P.: Compt. rend. **169,** 1095 (1919).
(100) JOLIBOIS, P.: Compt. rend. **169,** 1161 (1919).
(101) JOLIBOIS, P., AND MAZE-SENCIER, J.: Compt. rend. **181,** 36 (1925).
(102) JOLIBOIS, P., AND CLOUTIER, L.: Compt. rend. **195,** 1027 (1932).
(103) DE JONG, W. F.: Rec. trav. chim. **45,** 445 (1926).
(104) KAYSSER, A.: Chem.-Ztg. **44,** 826 (1920).
(105) KICK, C. H., BETHKE, R. M., AND EDGINGTON, B. H.: J. Agr. Research **46,** 1023 (1933).
(106) KICK, C. H., BETHKE, R. M., AND RECORD, P. R.: Poultry Sci. **12,** 382 (1933).
(107) KLEMENT, R. Z. physiol. Chem. **184,** 132 (1929).
(108) KLEMENT, R. Z. physiol. Chem. **185,** 237 (1929).
(109) KLEMENT, R. Z. physiol. Chem. **196,** 140 (1931).
(110) KLEMENT, R. AND TRÖMEL, G.: Z. physiol. Chem. **213,** 263 (1932).
(111) KLEMENT, R. AND TRÖMEL, G.: Klin. Wochschr. **12,** 292 (1933).
(112) KLEMENT, R. Z. physiol. Chem. **229,** 22 (1934); **235,** 272 (1935).
(113) KLEMENT, R. Ber. **68B,** 2012 (1935).
(114) KLEMENT, R. Naturwissenschaften **21,** 662 (1933).
(115) KLEMENT, R. Z. anorg. allgem. Chem. **228,** 232 (1936).
(116) KLEMENT, R. Z. anorg. allgem. Chem. **237,** 161 (1938).
(117) KLEMENT, R. Naturwissenschaften **26,** 145 (1938).
(118) KLEMENT, R. Ber. **70B,** 468 (1937).
(119) KÖRBER, F., AND TRÖMEL, G.: Z. Elektrochem. **38,** 578 (1932).
(120) KÖRBER, F., AND TRÖMEL, G.: Archiv. Eisenhüttenw. **7,** 7 (1933).
(121) KRAMER, B., AND SHEAR, M. J.: J. Biol. Chem. **79,** 147 (1928).
(122) KRAMER, B., SHEAR, M. J., AND MCKENZIE, M. R.: J. Biol. Chem. **82,** 555 (1929).
(123) KRAMER, B., SHEAR, M. J., AND SIEGAL, J.: J. Biol. Chem. **91,** 271, 723 (1931).
(124) KROLL, V. A.: J. Iron Steel Inst. (London) **84,** 126 (1911); Chem. Trade J. **49,** 367, 395, 421 (1912).

(125) Kyle, H. M.: Z. mikroskop.-anatom. Forsch. **9**, 347 (1927).
(126) Larson, H. W. E.: Ind. Eng. Chem., Anal. Ed. **7**, 401 (1935).
(127) LeFevre, M. L., and Bale, W. F.: Ind. Eng. Chem., Anal. Ed. **10**, 156 (1938).
(128) Logan, M. A.: J. Biol. Chem. **110**, 375 (1935).
(129) Logan, M. A., and Taylor, H. L.: J. Biol. Chem. **119**, 293 (1937).
(130) Logan, M. A., and Taylor, H. L.: J. Biol. Chem. **125**, 377 (1938).
(131) Logan, M. A., and Taylor, H. L.: J. Biol. Chem. **125**, 391 (1938).
(132) Lorah, J. R., Tartar, H. V., and Wood, L.: J. Am. Chem. Soc. **51**, 1097 (1929).
(133) MacIntire, W. H., Hardin, L. J., and Oldham, F. D.: Ind. Eng. Chem. **28**, 48 (1936).
(134) MacIntire, W. H., Hardin, L. J., Oldham, F. D., and Hammond, J. W.: Ind. Eng. Chem. **29**, 758 (1937).
(135) MacIntire, W. H., Hardin, L. J., and Oldham, F. D.: Ind. Eng. Chem. **29**, 224 (1937).
(136) MacIntire, W. H., and Hammond, J. W.: Ind. Eng. Chem. **30**, 160 (1938).
(137) Marek, J., Wellmann, O., and Urbányi, L.: Z. physiol. Chem. **226**, 3 (1934).
(138) Marek, J., Wellmann, O., and Urbányi, L.: Z. physiol. Chem. **229**, 24 (1934).
(139) Marek, J., Wellmann, O., and Urbányi, L.: Z. physiol. Chem. **234**, 165 (1935).
(140) Marek, J., Wellmann, O., and Urbányi, L.: Math. naturw. Anz. ungar. Akad. Wiss. **52**, 734 (1935).
(141) Marek, J., Wellmann, O., and Urbányi, L.: Z. physiol. Chem. **240**, 208 (1936).
(142) Marshall, H. L., Reynolds, D. S., Jacob, K. D., and Rader, L. F., Jr.: Ind. Eng. Chem. **27**, 205 (1935).
(143) Marshall, H. L., Reynolds, D. S., Jacob, K. D., and Tremearne, T. H.: Ind. Eng. Chem. **29**, 1294 (1937).
(144) McConnell, D.: Am. Mineral. **22**, 977 (1937).
(145) McConnell, D.: Am. Mineral. **23**, 1 (1938).
(146) McConnell, D.: Am. J. Sci. **36**, 296 (1938).
(147) Mehmel, M.: Z. Krist. **75**, 323 (1930); Z. physik. Chem. **B15**, 223 (1931).
(148) Messerschmitt, A.: Angew. Chem. **51**, 197 (1938).
(149) Miller, W. L., and Kenrick, F. B.: J. Phys. Chem. **7**, 259 (1903).
(150) Möller, H., and Trömel, G.: Naturwissenschaften **21**, 346 (1933).
(151) Möller, H., and Trömel, G.: Naturwissenschaften **24**, 377 (1936).
(152) Morgulis, S., with Janecek, E.: J. Biol. Chem. **93**, 455 (1931).
(153) Mückenberger, C.: Z. anorg. allgem. Chem. **169**, 81 (1928).
(154) Murray, M. M.: Biochem. J. **30**, 1567 (1936).
(155) Nacken, R.: Centr. Mineral. Geol. **1912**, 545.
(156) Nelson, O. A., and Haring, M. M.: J. Am. Chem. Soc. **59**, 2216 (1937).
(157) Nagelschmidt, G.: J. Chem. Soc. **1937**, 865.
(158) Neal, W. M., Palmer, L. S., Eckles, C. H., and Gullickson, T. W.: J. Agr. Research **42**, 115 (1931).
(159) Neal, W. M., and Palmer, L. S.: J. Agr. Research **42**, 107 (1931).
(160) Newton, R. H.: Chem. & Met. Eng. **45**, 374 (1938).
(161) Nielson, O.: Ferrum **10**, 97 (1912).
(162) Oswald, H.: Svenska Mosskulturföreningens Tid. **46**, 141 (1932); Phosphorsäure **3**, 507 (1933).

(163) Pearce, G. W., and Norton, L. B.: J. Am. Chem. Soc. **58**, 1104 (1936).
(164) Pearce, G. W., and Avens, A. W.: J. Am. Chem. Soc. **59**, 1258 (1937).
(165) Reynolds, D. S., Jacob, K. D., and Hill, W. L.: Ind. Eng. Chem. **21**, 1253 (1929).
(166) Reynolds, D. S., Jacob, K. D., and Rader, L. F., Jr.: Ind. Eng. Chem. **26**, 406 (1934).
(167) Reynolds, D. S., Jacob, K. D., Marshall, H. L., and Rader, L. F., Jr.: Ind. Eng. Chem. **27**, 87 (1935).
(168) Reynolds, D. S., Marshall, H. L., Jacob, K. D., and Rader, L. F., Jr.: Ind. Eng. Chem. **28**, 678 (1936).
(169) Reynolds, L., Corrigan, K. E., Hayden, H. S., Macy, I. G., and Hunocher, H. A.: Am. J. Roentgenol. Radium Therapy **39**, 103 (1938).
(170) Reynolds, L., Hayden, H. S., and Corrigan, K. E.: Am. J. Roentgenol. Radium Therapy **39**, 286 (1938).
(171) Rindell, A.: Helsingfors, 1899.
(172) Rindell, A.: Compt. rend. **134**, 112 (1902).
(173) Robison, R.: Ann. Rev. Biochem. **5**, 181 (1936).
(174) Rogers, A. F.: Mineral. Mag. **17**, 155 (1914).
(175) Rogers, A. F.: J. Geology **25**, 515 (1917).
(176) Rogers, A. F.: Am. J. Sci. **3**, 269 (1922).
(177) Rogers, A. F.: Bull. Geol. Soc. Am. **35**, 535 (1924).
(178) Roseberry, H. H., Hastings, A. B., and Morse, J. K.: J. Biol. Chem. **90**, 395 (1931).
(179) Ross, W. H., and Jacob, K. D.: J. Assoc. Official Agr. Chem. **20**, 231 (1937).
(180) Russell, E. J.: Trans. Faraday Soc. **16**, 263 (1921).
(181) St. Náray-Szábo: Z. Krist. **75**, 387 (1930).
(182) St. Jacques, C.: Ind. Eng. Chem., News Ed. **15**, 29 (1937).
(183) Sanfourche, A., and Focet, B.: Bull. soc. chim. **53**, 974 (1933).
(184) Sanfourche, A., and Henry, J.: Bull. soc. chim. **53**, 1210 (1933); Compt. rend. **194**, 1940 (1932).
(185) Sanfourche, A.: Bull. soc. chim. **53**, 951 (1933).
(186) Sanfourche, A.: Bull. soc. chim. **53**, 1507 (1933).
(187) Sanfourche, A.: Compt. rend. **196**, 935 (1933).
(188) Saupe, E.: Kolloid-Z. **69**, 357 (1934).
(189) Schaller, W. T.: J. Wash. Acad. Sci. **1**, 151 (1911).
(190) Schaller, W. T.: U. S. Geol. Survey, Bull. 509, 115 pp. (1912).
(191) Schleede, A., Schmidt, W., and Kindt, H.: Z. Elektrochem. **38**, 633 (1932).
(192) Schmidt, C. L. A., and Greenberg, D. M.: Physiol. Rev. **15**, 297 (1935).
(193) Schneiderhöhn, H.: Mitt. Kaiser-Wilhelm Inst. Eisenforsch. Düsseldorf **13**, 109 (1931).
(194) Schneiderhöhn, H.: Arch. Eisenhüttenw. **5**, 9 (1931).
(195) Schwarz, R., Eden, R., and Herrmann, E.: Biochem. Z. **149**, 100 (1924).
(196) Sendroy, J., Jr., and Hastings, A. B.: J. Biol. Chem. **71**, 783 (1927).
(197) Sendroy, J., Jr., and Hastings, A. B.: J. Biol. Chem. **71**, 797 (1927).
(198) Serrales, J. J., Jr.: Soil Sci. **44**, 175 (1937).
(199) Sharpless, G. R., and McCollum, E. V.: J. Nutrition **6**, 163 (1933).
(200) Shear, M. J., and Kramer, B.: J. Biol. Chem. **79**, 105 (1928).
(201) Shear, M. J., and Kramer, B.: J. Biol. Chem. **79**, 125 (1928).
(202) Shear, M. J., Washburn, M., and Kramer, B.: J. Biol. Chem. **83**, 697 (1929).
(203) Shear, M. J., and Kramer, B.: J. Biol. Chem. **86**, 677 (1930).

(204) Shohl, A. T., Brown, H. B., Chapman, E. E., Rose, C. S., and Saurwein, E. M.: J. Nutrition **6**, 271 (1933).
(205) Silberstein, L.: Compt. rend. **200**, 421 (1935).
(206) Smith, C. M., and Murray, C. W.: Ind. Eng. Chem. **23**, 207 (1931).
(207) Smith, G. F. H., and Prior, G. T.: Mineral. Mag. **16**, 84 (1911).
(208) Smith, H. V., and Smith, M. C.: Water Works Eng. **90**, 1600 (1937).
(209) Statsmann, L.: Biochem. Z. **217**, 395 (1930).
(210) Tartar, H. V., Wood, L., and Hiner, E.: J. Am. Chem. Soc. **46**, 809 (1924).
(211) Taylor, N. W., and Sheard, C.: J. Biol. Chem. **81**, 479 (1929).
(212) Thewlis, J.: Brit. J. Radiol. **5**, 353 (1932).
(213) Thewlis, J.: Brit. J. Radiol. **9**, 300 (1936).
(214) Thewlis, J.: Nature **137**, 828 (1936).
(215) Travers, A., and Bauer: Compt. rend. **196**, 1802 (1933); Helv. Chim. Acta **16**, 918 (1933).
(216) Trebitsch, F.: Biochem. Z. **191**, 234 (1927).
(217) Trömel, G.: Phosphorsäure **2**, 116 (1932).
(218) Trömel, G.: Mitt. Kaiser-Wilhelm Inst. Eisenforsch. Düsseldorf **14**, 25 (1932).
(219) Trömel, G.: Z. physik. Chem. **A158**, 422 (1932).
(220) Trömel, G., and Möller, H.: Z. anorg. allgem. Chem. **206**, 227 (1932).
(221) Trömel, G.: Phosphorsäure **3**, 129 (1933).
(222) Walden, G. H., Jr., and Cohen, M. U.: J. Am. Chem. Soc. **57**, 2591 (1935).
(223) Warington, R.: J. Chem. Soc. **26**, 983 (1873).
(224) Wendt, G. L., and Clarke, A. H.: J. Am. Chem. Soc. **45**, 881 (1923).
(225) Werner, A.: Ber. **40**, 4441 (1907).
(226) Whittaker, C. W., Adams, J. R., and Jacob, K. D.: Ind. Eng. Chem. **29**, 1144 (1937).
(227) Wilhelmj, A.: Z. Pflanzenernähr. Düngung Bodenk. **19A**, 129 (1931); Phosphorsäure **2**, 494 (1932).
(228) Withers, W. A., and Feild, A. L.: J. Am. Chem. Soc. **37**, 1091 (1915).
(229) U. S. patents: 1,902,832 (March 28, 1933); 2,044,774 (June 23, 1936); 2,070,697 (Feb. 16, 1937); 2,093,176 (Sept. 14, 1937); 2,100,843 (Nov. 30, 1937); 2,121,776 (June 28, 1938).
German patents: 662,913 (July 26, 1938); 663,460 (Aug. 6, 1938).
French patents: 766,874 (July 5, 1934); 775,457 (Dec. 28, 1934); 799,190 (June 8, 1936); 800,726 (July 17, 1936); 809,554 (March 6, 1937).
British patents: 418,788 (Oct. 31, 1934); 463,849 (April 7, 1937); 467,126 (June 11, 1937); 469,833 (Aug. 4, 1937); 473,313 (Oct. 11, 1937).
(230) U. S. patent: 2,126,793 (Aug. 16, 1938).

A SUMMARY OF THE REACTIONS OF ALDEHYDES WITH AMINES

MURRAY M. SPRUNG

Research Laboratory, General Electric Company, Schenectady, New York

Received August, 1939

CONTENTS

I. INTRODUCTION

Most reactions of aldehydes involve the carbonyl oxygen and are, in the first instance, addition reactions. Primary and secondary amines are especially suited to the study of the addition reactions of aldehydes, since they contain one or two easily detachable ("active") hydrogen atoms.

The first references to the reactions of aldehydes and amines appeared nearly a century ago, and a voluminous literature has grown up in this field. The present paper will attempt to summarize and systematize the literature from about 1880 to the present date (approximately June, 1939).

The general features of the problem may be stated in very simple terms. In most cases the first recognizable step is addition of the amine to the carbonyl group of the aldehyde, in the sense of the first phase of the aldol condensation.

$$RCHO + R'NH_2 \rightarrow RCHOHNHR'$$

In the majority of cases, the α-hydroxy amine which results is not capable of isolation; it reacts further, in one of several ways. It may, for example, lose water, to give an imine (Schiff base); this corresponds to the "crotonaldehyde stage" of the aldol condensation.[1]

$$RCHOHNHR' \rightarrow RCH{=}NR' + H_2O$$

Alternately, the hydroxy amine may react further with one or the other of the reagents or with itself; or the Schiff base may be involved in the various subsequent reactions which are encountered, and which often lead to the formation of complex reaction products.

The attempt will be made to indicate as many as possible of the types of "follow" reactions which occur. Since the nature of the aldehyde and of the amine is the determining factor, even more frequently than the experimental conditions, it has been found convenient to subdivide the material under discussion so as to accord with a classification of the reactants into "aliphatic–aliphatic," "aliphatic–aromatic," and "aromatic–aromatic" categories. This order is indicated in the table of contents.

Since tertiary amines bear no "active" hydrogen, they seldom enter into reaction with aldehydes, unless the amino nitrogen is attached directly to an aromatic nucleus. Tertiary amines will therefore not be considered in the present review.

II. REACTION OF ALIPHATIC ALDEHYDES WITH AMMONIA AND WITH ALIPHATIC AMINES

A. *Formaldehyde and ammonia*

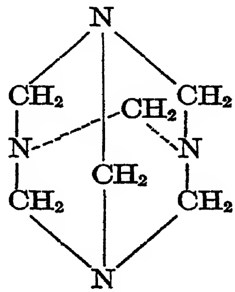
The reaction between ammonia and formaldehyde, curiously enough, occupies an almost unique position in the present discussion. When

[1] We shall refer to this type of imine in the discussion which is to follow as a Schiff base, although this term has often been restricted to the imines formed from aromatic aldehydes and amines. We shall not use the term "azomethine", which has frequently been used to designate the essential C═N linkage involved here.

ammonia and formaldehyde are mixed in aqueous solution, hexamethylenetetramine is formed (as was first observed seventy-eight years ago by Butlerow (11)), according to the following equation:

$$6H_2O + 4NH_3 \rightarrow (CH_2)_6N_4 + 6H_2O$$

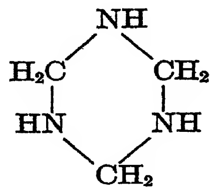
Hexamethylenetetramine is a well-defined, crystalline solid, the crystal structure of which is known in detail (27, 159). All six carbon atoms are stereochemically equivalent, as are the four nitrogen atoms. The lines joining carbon atoms and nitrogen atoms in the molecular model form a regular octahedron and a regular tetrahedron, respectively. This structure possesses a very high degree of symmetry. The generally accepted structure of hexamethylenetetramine,



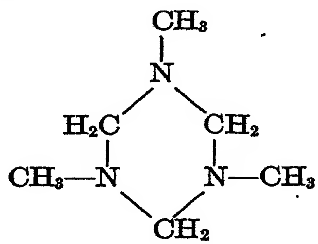
follows from the crystallographic data, from its molecular weight, and from the reactions involved in its synthesis and degradation (19, 34, 120). The construction of a model easily convinces one that this structure involves four sterically equivalent, six-membered, virtually strainless rings, and that the molecule is completely symmetrical. An alternate formula for hexamethylenetetramine, however, has frequently been proposed, and is still presented in some texts, as follows: $N(CH_2N{=}CH_2)_3$. The latter formula is said to explain more readily the easy regeneration of formaldehyde and ammonia from hexamethylenetetramine, as well as the fact that the compound acts as a monoacid base, thereby indicating that one nitrogen atom may be different from the other three.

Despite the fact that this stable end product is the only one isolated in the reaction between ammonia and formaldehyde, there is considerable evidence that a number of simpler intermediate substances are formed, and that aqueous solutions of hexamethylenetetramine break down through a succession of stages involving these same intermediates. Thus it is possible that mono-, di-, and tri-methylolamines, NH_2CH_2OH, $NH(CH_2OH)_2$, and $N(CH_2OH)_3$, are all formed, by a series of successive addition reactions. As a matter of fact, the sulfite esters of these substances have been prepared by the interaction of formaldehyde and ammonia in the presence of sodium bisulfite (17, 103). Another postulated

intermediate in the reaction is methylenediamine, $NH_2CH_2NH_2$, the nitrogen analog of methylene glycol. Salts of this substance have been obtained by treating methylene bisformamide, $HCONHCH_2NHCHO$, with strong acids (68). The formation of the six-membered ring compound, trimethylenetriamine,



has also been postulated (34), by analogy with the known formation of trimethyltrimethylenetriamine,



from methylamine and formaldehyde, and the formation of trioxymethylene by polymerization of formaldehyde. It is likely that trimethylenetriamine is formed, but that the presence of the three active hydrogen atoms in this substance enables the condensation with formaldehyde and ammonia to continue until the completely saturated and completely symmetrical structure represented by hexamethylenetetramine ultimately results.

In the presence of acids, the reaction between ammonia and formaldehyde takes an entirely different course. When formaldehyde is heated with ammonium sulfate or ammonium chloride, the products are mono-, di-, and tri-methylamines, formic acid, and carbon dioxide (98, 153). This reaction apparently involves simultaneous oxidation and reduction of the formaldehyde. It is more likely, however, that methylolamines are first formed, and are then reduced by excess formaldehyde to methylamines, e.g.:

$$CH_2O + NH_3 \xrightarrow{\text{addition}} HOCH_2NH_2 \xrightarrow[\text{reduction}]{CH_2O} CH_3NH_2 \xrightarrow[\text{addition}]{CH_2O}$$

$$CH_3NHCH_2OH \xrightarrow[\text{reduction}]{CH_2O} CH_3NHCH_3, \text{ etc.}$$

$$(\text{Oxidation of } CH_2O \rightarrow HCOOH,\ CO_2,\ \text{and } H_2O)$$

B. Aliphatic aldehydes and ammonia

When aliphatic aldehydes of low molecular weight are treated in ether solution with gaseous ammonia, or when their aqueous or alcoholic solutions are treated with strong ammonia, crystalline solids are obtained which correspond in composition to simple addition compounds. These are the "aldehyde ammonias." They are relatively unstable substances, for on standing in moist air, or on treatment with dilute acids, they readily revert into their constituents.

$$RCHO + NH_3 \rightleftharpoons RCHOHNH_2$$

Furthermore, on standing over concentrated sulfuric acid or on heating in a vacuum they lose water and give imines, usually as trimers.

$$RCHOHNH_2 \rightarrow H_2O + RCH{=}NH \rightarrow (\underset{|}{R}CH{-}NH{-})_3$$

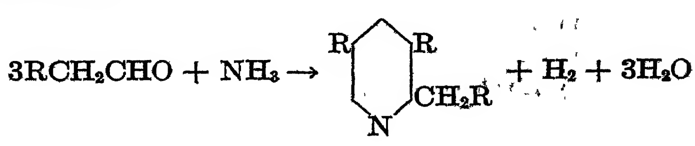
On long standing the aldehyde ammonias are changed to amorphous materials, presumably by dehydration and polymerization.

The aldehyde ammonias themselves do not in general correspond to the simple monomeric formula pictured above. Thus, acetaldehyde ammonia has a molecular weight in solution which corresponds to something between a dimeric and trimeric structure (2, 20). In the vapor phase it is partially dissociated into monomeric ethylidenimine, $CH_3CH{=}NH$. At 260°C. and 30 mm. the dissociation is complete (117). If the vapors are then suddenly cooled, the liquid ethylidenimine monomer may be obtained, but this is slowly converted to the trimer on standing. The latter has been shown to have the usual six-membered ring structure (23).

The aldehyde ammonia derived from chloral is monomeric in acetic acid, dimeric in benzene, and dimeric or higher in ethylene dibromide (21). The cyclic trimer has been reported to exist in two isomeric modifications, one melting at 105–6°C. and the other at 150–5°C. (3).

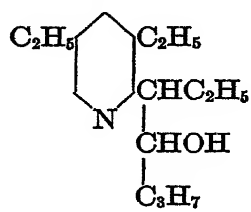
1. The Tschitschibabin reaction

An interesting and important reaction was discovered in 1905 by Tschitschibabin. It consists in passing the vapors of an aliphatic aldehyde and ammonia over alumina at 300–400°C. Pyridine derivatives are thus formed (61, 117, 122, 123, 124, 125). The reaction is generally formulated as follows:



Thus with acetaldehyde the main product is α-picoline; with propionaldehyde it is 2-ethyl-3,5-dimethylpyridine; with butyraldehyde it is 2-propyl-3,5-diethylpyridine; with isovaleraldehyde one obtains 2-isobutyl-3,5-diisopropylpyridine; and with heptaldehyde one obtains 2-hexyl-3,5-diamylpyridine. However, more highly alkylated pyridines are also observed as by-products, owing presumably to the fact that the aliphatic aldehyde aldolizes during the course of the reaction, and the aldolization product then reacts with ammonia. Thus, with acetaldehyde the formation of methylethylpyridine and of β-collidine was observed in addition to α- and γ-picoline.

The Tschitschibabin reaction may also be effected by heating aldehyde ammonias. In the case of butyraldehyde ammonia, the expected 2-propyl-3,5-diethylpyridine is obtained, along with a substantial yield of a more complex substance, to which the following structure has been assigned and which is formed from the former substance by an aldol condensation with an additional molecule of butyraldehyde (61):

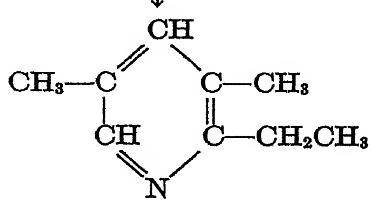


Strain (117) has represented the course of the Tschitschibabin reaction (in the case of propionaldehyde) as follows:

$$CH_3CH_2CHO + NH_3 \rightarrow CH_3CH_2CH{=}NH + H_2O$$

$$3CH_3CH_2CH{=}NH \rightarrow CH_3CH_2\underset{NH_2}{\underset{|}{C}H}-\underset{CH_3}{\underset{|}{C}H}-\underset{NH_2}{\underset{|}{C}H}-\underset{CH_3}{\underset{|}{C}H}CH{=}NH$$

$$\downarrow$$



$$+ 2NH_3 + H_2$$

Thus the first step is presumably the formation of an imine, the second step is an aldol-like trimerization of the imine, and the third step is ring-closure with concomitant loss of ammonia and hydrogen.

The case of crotonaldehyde is somewhat distinctive. A compound of

uncertain structure is formed, which has been given the name "tricrotylidenetetramine" and the supposed formula $C_{12}H_{24}N_4$. Delépine (24) has suggested that this reaction be represented as follows:

$$3CH_3CH{=}CHCHO + 4NH_3 \rightarrow \begin{array}{l} \qquad\qquad\qquad\quad NH_2 \\ CH_3CH{=}CHCH\Big\langle \\ \qquad\qquad\qquad\quad NH \\ CH_3CH{=}CHCH\Big\langle \\ \qquad\qquad\qquad\quad NH \\ CH_3CH{=}CHCH\Big\langle \\ \qquad\qquad\qquad\quad NH_2 \end{array}$$

If this representation is accurate, tricrotylidenetetramine is closely related structurally to the well-known and technically valuable polyethylene polyamines.

C. Aliphatic aldehydes and aliphatic amines

1. Aliphatic aldehydes and aliphatic monoamines

It has previously been indicated that aliphatic amines can be methylated by heating with formaldehyde in the presence of acids. However, if the reaction is carried out at room temperature and in the presence of alkali, it takes another course; addition results and the corresponding methylol derivatives can be isolated. Both primary and secondary aliphatic amines react in this way (64).

$$RNH_2 + CH_2O \xrightarrow{OH^-} RNHCH_2OH$$

$$R_2NH + CH_2O \xrightarrow{OH^-} R_2NCH_2OH$$

The methylolamines derived from methyl-, ethyl-, propyl-, isobutyl-, isoamyl-, and benzyl-amines and from piperidine are colorless liquids. They can be condensed with a second molecule of amine to give bis(alkylamino)- or bis(dialkylamino)-methanes.

$$RNHCH_2OH + R'NH_2 \rightarrow RNHCH_2NHR' + H_2O$$

$$R_2NCH_2OH + R_2'NH \rightarrow R_2NCH_2NR_2' + H_2O$$

The bis(alkylamino)methanes can also be formed directly from the amine and formaldehyde (70).

When distilled over solid potassium hydroxide, the methylolalkylamines lose water to give the corresponding alkylmethylenimines (or Schiff bases), usually as the cyclic trimers (14, 41, 47, 63).

$$3RNHCH_2OH \rightarrow (RN{-}CH_2{-})_3 + 3H_2O$$

The Schiff bases derived from the higher aliphatic aldehydes are sometimes sufficiently stable to be isolated as the monomers. Examples are ethylideneëthylimine, $C_2H_5N{=}CHCH_3$, b.p. 48°C. (65), and propylidenepropylimine, $CH_3CH_2CH_2N{=}CHCH_2CH_3$, b.p. 102°C. (16).

Secondary aliphatic amines may be caused to condense in a similar fashion with higher aliphatic aldehydes, by treating them with solid potassium carbonate. The products correspond to the bis(dialkylamino)-methanes described above (85). These ditertiary amines usually lose one molecule of amine on distillation, giving rise to α,β-unsaturated amines (enamines), e.g.,

$$\underset{\text{Piperidine}}{2C_5H_{10}NH} + CH_3CH_2CH_2CHO \xrightarrow{K_2CO_3}$$

$$(C_5H_{10}N)_2CHCH_2CH_2CH_3 + H_2O$$

$$\downarrow$$

$$C_5H_{10}NCH{=}CHCH_2CH_3 + C_5H_{10}NH$$

Diethylamine and methylaniline behave similarly. Among the aldehydes which react readily are propionaldehyde, butyraldehyde, acetaldehyde, and hydrocinnamaldehyde. In the case of phenylacetaldehyde and dibenzylamine, the unsaturated amine, dibenzylstyrylamine, is formed spontaneously. This change may be represented stoichiometrically as follows:

$$C_6H_5CH_2CHO + (C_6H_5CH_2)_2NH \rightarrow (C_6H_5CH_2)_2NCH{=}CHC_6H_5 + H_2O$$

It is not known whether aldehyde ammonias are formed as intermediates, or whether the amine condenses directly with the aldehyde through the enol form of the latter.

With α,β-unsaturated aldehydes, such as acrolein, crotonaldehyde, cinnamaldehyde, and citral, crystalline condensation products are obtained only at temperatures as low as $-10°$ to $-20°$C.; at higher temperatures, resins or complex aldol-like condensation products result (86). The crystalline products thus obtained may be represented as resulting from a 1,4-addition to the conjugated system, followed by condensation with a second molecule of amine.

$$R'CH{=}CHCHO + R_2NH \longrightarrow \underset{\substack{|\\ NR_2}}{R'CH}CH{=}CHOH \xrightarrow{R_2NH}$$

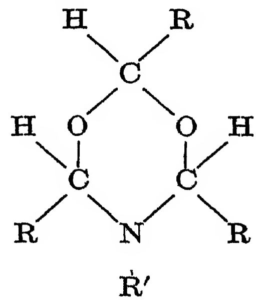
$$\underset{\substack{|\\ NR_2}}{R'CH}CH{=}CHNR_2 + H_2O$$

The unsaturated ditertiary amines decompose on distillation into a molecule of secondary amine and a molecule of a dialkylaminoallene.

$$R'\underset{\underset{NR_2}{|}}{C}HCH{=}CHNR_2 \longrightarrow R_2NH + R'CH{=}C{=}CHNR_2$$

A somewhat unusual type of reaction has been observed to occur when ethylamine or allylamine reacts with formaldehyde or acetaldehyde at ice-bath temperature. Cyclic compounds are produced which correspond to copolymerization products of two molecules of the aldehyde with one molecule of the monomeric alkylidenimine (4).

$3RCHO + R'NH_2$



$+ H_2O$

The mechanism of this condensation has not been elucidated.

2. Aliphatic aldehydes and aliphatic diamines and polyamines

Ethylenediamine, trimethylenediamine, tetramethylenediamine, and pentamethylenediamine react with formaldehyde in alkaline solution to give products whose exact structure has not been determined, but which correspond to condensation products of 2 moles of aldehyde with 1 mole of diamine, according to the following equation (6, 7):

$$NH_2(CH_2)_nNH_2 + 2CH_2O \rightarrow —CH_2—\underset{|}{N}(—CH_2)_n—\underset{|}{N}—CH_2— + 2H_2O$$

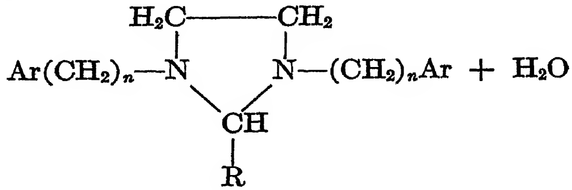
The product derived from ethylenediamine is a crystalline solid, having a molecular weight in benzene twice that of the monomeric unit shown above. Those derived from the three- and four-carbon diamines are liquids, boiling between 200° and 300°C., and are of indefinite molecular weight. That derived from the five-carbon diamine is a high-melting solid, which suggests that it may be polymeric.

The reaction of isobutyraldehyde on ethylenediamine gives rise to a similar condensation product, which has been given the structure of diisobutylideneëthylenediamine, $(CH_3)_2CHCH{=}NCH_2CH_2N{=}CHCH(CH_3)_2$ (69).

1,2-Bis(benzylamino)ethane, 1,2-bis(p-methylbenzylamino)ethane, 1,2-

bis(β-phenylethylamino)ethane, 1,2-bis(isobutylamino)ethane, and 1,2-bis(furfurylamino)ethane (all of which may be considered as aliphatic diamines, since the amino groups are not attached directly to an aromatic nucleus) react with aldehydes to give tetrahydroimidazoles (78, 102). The general reaction involved may be represented by the following equation:

$$Ar(CH_2)_nNHCH_2CH_2NH(CH_2)_nAr + RCHO \longrightarrow$$



The aliphatic aldehydes which enter into this reaction include formaldehyde, acetaldehyde, propionaldehyde, butyraldehyde, isobutyraldehyde, valeraldehyde, and caprylaldehyde. Apparently the closure of the five-membered heterocyclic ring is greatly facilitated when the nitrogen atoms which enter into the reaction are secondary.

III. REACTION OF AROMATIC ALDEHYDES WITH AMMONIA AND WITH ALIPHATIC AMINES

A. Aromatic aldehydes and ammonia

Aromatic aldehydes in general react with aqueous or alcoholic ammonia at room temperature to give the so-called "hydroamides." These are high-melting, crystalline substances formed by combination of three molecules of the aldehyde and two molecules of ammonia, as follows:

$$3ArCHO + 2NH_3 \rightarrow ArCH(N{=}CHAr)_2 + 3H_2O$$

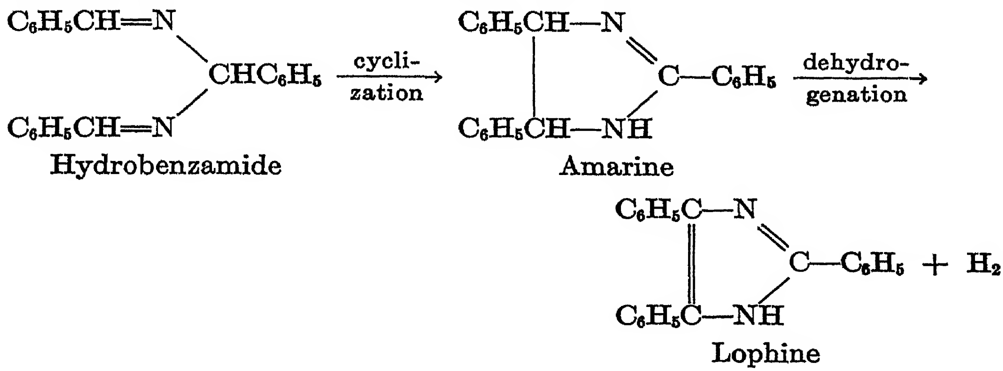
In the cases of benzaldehyde and *p*-methylbenzaldehyde, an intermediate, crystalline addition compound, consisting of two molecules of the aldehyde and one molecule of ammonia, has been isolated by Francis by carrying out the reaction at low temperatures (48). When warmed, the addition compound breaks up into the hydroamide, aldehyde, and water. The mechanism of the reaction may therefore be represented as follows:

$$2ArCHO + NH_3 \rightarrow ArCHOHNHCHOHAr$$

$$2ArCHOHNHCHOHAr \rightarrow (ArCH{=}N)_2CHAr + ArCHO + 3H_2O$$

Certain hydroamides, when heated in the neighborhood of 130°C. for a few hours, are cyclized to the corresponding 2,4,5-triaryldihydro-

imidazoles (9, 22, 25, 45, 157). Heating to a still higher temperature results in dehydrogenation to the corresponding triarylimidazole. These changes are illustrated, for the case of hydrobenzamide, as follows:



Exactly similar relationships obtain in the cases of hydroanisamide and hydrofurfuramide. In the case of cinnamaldehyde, the dihydroimidazole is formed directly upon treatment of the aldehyde with aqueous ammonia. Hydrosalicylamide, on the other hand, does not readily rearrange to the amarine structure. *p*-Hydroxybenzaldehyde does not give a stable hydroamide.

1. Reaction kinetics

The kinetics of the aromatic aldehyde–ammonia reaction were investigated by Dobler (30), who found them to be in accord with a bimolecular mechanism in the case of ammonia and benzaldehyde. There was an induction period which extended for about 100 min. at 20°C. A marked falling off of the reaction velocity was noted at about 48 per cent reaction, an observation which was attributed to the increased influence of the reverse reaction. Francis' mechanism (outlined above) was used as a working hypothesis; it was assumed that the first step—formation of the bis(phenylmethylol)amine—was a fast reaction, and that the rate-controlling step was the subsequent bimolecular decomposition of this intermediate into hydrobenzamide and benzaldehyde. The assumption could then be made that the induction period was due to the building up of the bis(phenylmethylol)amine. However, during this period the reaction should follow a trimolecular course (if Francis' mechanism is strictly valid), which was not observed to be the case. The effect of increasing the temperature was found to be a marked decrease in the concentration of hydrobenzamide present at equilibrium; above about 57°C. very little of this product would be formed.

Reaction velocity constants were determined for a number of other

aldehydes and found to increase in the following order: anisaldehyde, *p*-nitrobenzaldehyde, *o*-nitrobenzaldehyde, benzaldehyde, *m*-xylylaldehyde, *m*-nitrobenzaldehyde, *p*-tolualdehyde, *p*-chlorobenzaldehyde, and cinnamaldehyde. There was a sevenfold spread in rates from anisaldehyde to *p*-chlorobenzaldehyde, while cinnamaldehyde reacted with an apparent speed twice that of *p*-chlorobenzaldehyde. Dobler did not take into account the fact that cinnamaldehyde goes directly to the dihydroimidazole structure on treatment with aqueous ammonia.

B. Aromatic aldehydes and aliphatic amines

The products most easily isolated when aromatic aldehydes are caused to condense with aliphatic amines are the Schiff bases. The "aldol" is probably first formed, but is rarely capable of isolation. Table 1 shows examples of this type of Schiff base.

Aliphatic diamines (and some polyamines) react readily with aromatic aldehydes. Diprimary amines usually give bis(arylidene)imines (87, 118, 126).

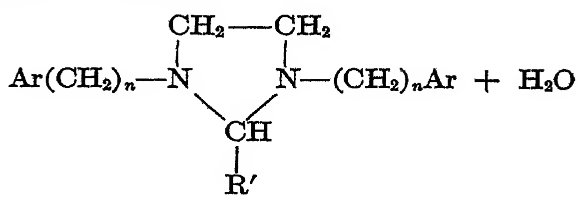
$$ArCHO + NH_2(CH_2)_nNH_2 \rightarrow ArCH{=}N(CH_2)_nN{=}CHAr + 2H_2O$$

The following are a few examples:

	MELTING POINT
	°C.
1,2-Bis(benzylidenimino)ethane	53–4
1,2-Bis(*p*-isopropylbenzylidenimino)ethane	63–4
1,2-Bis(cinnamylidenimino)ethane	109–10
1,2-Bis(*o*-hydroxybenzylidenimino)ethane	125–6
1,2-Bis(*o*-methoxybenzylidenimino)ethane	113
1,2-Bis(*p*-methoxybenzylidenimino)ethane	110–11

The reaction of aromatic aldehydes on aliphatic 1,2-disecondary amines readily leads to ring closure, with the formation of tetrahydroimidazoles (78, 102, 126). The general reaction involved is analogous to that which occurs with aliphatic aldehydes (page 306).

$$Ar(CH_2)_nNHCH_2CH_2NH(CH_2)_nAr + Ar'CHO \longrightarrow$$



Tetrahydroimidazoles have thus been prepared from 1,2-bis(benzyl-

amino)ethane, 1,2-bis(β-phenylethylamino)ethane, 1,2-bis(*p*-methoxybenzylamino)ethane, 1,2-bis(isobutylamino)ethane, and 1,2-bis(furfurylamino)ethane with the following aldehydes: benzaldehyde, *p*-tolualdehyde, *p*-methoxybenzaldehyde, *o*-, *m*-, and *p*-nitrobenzaldehydes, *o*-, *m*-, and *p*-chlorobenzaldehydes, salicylaldehyde, *p*-hydroxybenzaldehyde, *p*-hydroxy-*m*-methoxybenzaldehyde, piperonal, phenylacetaldehyde, furfural,

TABLE 1

Schiff bases formed by the condensation of aromatic aldehydes with aliphatic amines

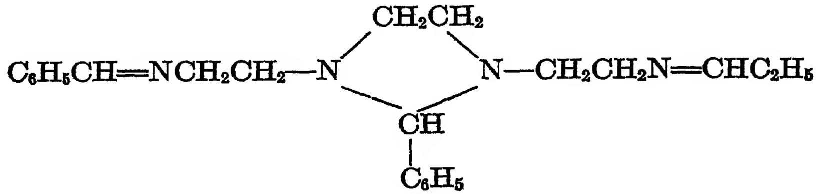
SCHIFF BASE	BOILING POINT	MELTING POINT	REFERENCE
	°C.	°C.	
Benzylidenemethylimine	180		(161)
Benzylideneëthylimine	195		(161)
Benzylidenepropylimine	208–10		(161)
Benzylideneisobutylimine	217–18 (35 mm.)		(161)
Benzylideneisoamylimine*			(107)
Benzylideneallylimine	96 (12 mm.)		(4)
Benzylidene-α-methyldecylimine	197–8 (17 mm.)		(99)
Benzylidene-β,β-diethoxyethylimine	156 (12 mm.)		(44)
Benzylidene-β,β-diethoxypropylimine	157 (11 mm.)		(158)
2,5-Dichlorobenzylidenemethylimine		52	(54)
o-Nitrobenzylidenemethylimine	145 (23 mm.)		(1)
p-Isopropylbenzylidenemethylimine	122 (14 mm.)		(112)
Cinnamylideneëthylimine	143–5 (20 mm.)		(1)
o-Hydroxybenzylidenemethylimine	229		(26)
o-Hydroxybenzylideneëthylimine	237		(26)
o-Hydroxybenzylidene-β-bromoethylimine		56–7	(53)
o-Hydroxybenzylidene-β,β-diethoxyethylimine	188 (15 mm.)		(62)
m-Hydroxybenzylidene-β,β-diethoxyethylimine		71	(52)
m-Methoxybenzylidene-β,β-diethoxyethylimine	191 (15 mm.)		(52)
m-Ethoxybenzylidene-β,β-diethoxyethylimine	220 (35 mm.)		(52)
p-Methoxybenzylidene-β,β-diethoxyethylimine	190 (12 mm.)		(62)

* A heavy oil.

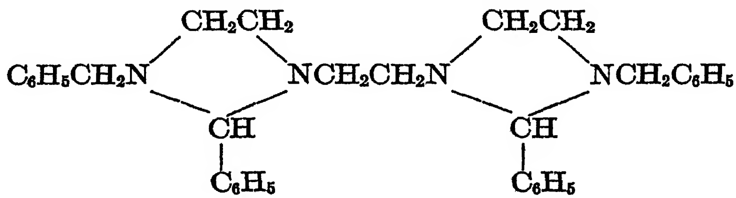
α-methylfurfural, α-methoxyfurfural, and α-hydroxymethylfurfural. Most of these tetrahydroimidazoles are colorless oils boiling in the neighborhood of 200°C. at about 20 mm. A few of them are crystalline solids.

van Alphen has shown that when the reaction is carried out with aliphatic 1,2-disecondary amines which also bear terminal primary amino groups, 3 moles of the aldehyde are required and the product is a bis(aryl-

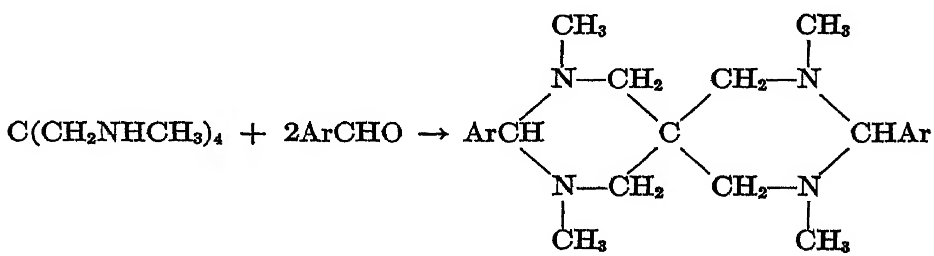
ideniminoalkyl)tetrahydroimidazole (126). Thus, triethylenetetramine and benzaldehyde give a compound of the following structure:



This substance was reduced with sodium and alcohol, and the benzylidene group hydrolyzed with acid. The resulting tetrasecondary amine was again treated with benzaldehyde, giving the following interesting compound, containing two tetrahydroimidazole rings:



van Alphen (127) also prepared a series of very interesting spiro compounds by causing aromatic aldehydes to react with tetrakis(methylaminomethyl) methane.



The aldehydes used in this series of syntheses were benzaldehyde, anisaldehyde, *p*-chlorobenzaldehyde, *p*-nitrobenzaldehyde, and piperonal. The spiro hexahydro-1,3-diazines are crystalline solids.

IV. REACTION OF ALIPHATIC ALDEHYDES WITH AROMATIC AMINES

A large part of the literature in the field of amine–aldehyde reactions is concerned with the action of aliphatic aldehydes (especially formaldehyde) on aromatic amines. To avoid confusion, it seems desirable at this point to consider some of the individual reactions separately.

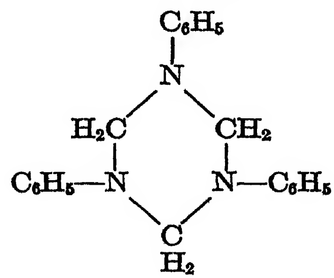
A. Formaldehyde

1. Formaldehyde and aromatic monoamines

(*a*) *Aniline.* The literature on this reaction goes back at least seventy years. Much confusion obtained for many years, and many contradictions have not yet been reconciled. An explanation lies in the fact that both crystalline and resinous substances are formed when aniline is treated with formaldehyde. At present we shall consider only the crystalline substances, and devote a separate section to the resins later.

When aniline is treated with formaldehyde in dilute, neutral aqueous solution, products of indefinite composition are obtained unless the conditions are carefully controlled. Under favorable conditions, a compound of the empirical composition C_7H_7N and having a melting point of 140°C. may be isolated (6, 70, 100, 101, 119, 136, 152).

It is now generally recognized that this substance, known as anhydroformaldehyde aniline, is a trimer of the hypothetical Schiff base, methyleneaniline, $C_6H_5N{=}CH_2$, and has a stable six-membered ring structure—namely, that of triphenyltrimethylenetriamine,



However, it is apparently capable of dissociating at higher temperatures, and shows a molecular weight in camphor which is only about twice that of the monomer (93). It is stated that the molecular weight of the vapor corresponds to that of the monomer (70).

Neither the initial "aldol" nor the Schiff base has ever been isolated. However, in alkaline solution one or the other of these must have a transitory existence, for the product which is then isolated is bis(phenylamino)-methane, $(C_6H_5NH)_2CH_2$, (m.p. 64–5°C.), which would result from the further reaction of either of these hypothetical substances with a second mole of aniline (35). Bis(phenylamino)methane is also sometimes found in neutral solution, along with anhydroformaldehyde aniline (37).

The reaction takes a much more complicated course in acid solution, and it is under such conditions that the technically useful aniline–formaldehyde resins are obtained. Among the intermediate substances which are formed in acid solution are the following: aminobenzylaniline,

$NH_2C_6H_4CH_2NHC_6H_5$; diaminodiphenylmethane, $NH_2C_6H_4CH_2C_6H_4NH_2$[2]; and anhydroaminobenzyl alcohol, $(—NHC_6H_4CH_2—)_x$ (141). More light will be thrown on some phases of this problem when the action of formaldehyde on *p*-toluidine and similar bases is discussed.

(*b*) *Toluidines.* In neutral solution, anhydroformaldehyde *o*-toluidine, $(o\text{-}CH_3C_6H_4NCH_2—)_x$, is formed from *o*-toluidine and formaldehyde (x is probably 3) (152). Under carefully controlled conditions, or in the presence of alkali, the product is bis(*o*-tolylamino)methane (m.p. 52°C.) (35). In acid solution products of indeterminate structure usually result (74, 95). If half a mole of formaldehyde is used in the presence of dilute acid, the chief product is 4,4′-diamino-3,3′-dimethyldiphenylmethane (m.p. 158°C.) (141).

In neutral solution, under ordinary circumstances, anhydroformaldehyde *m*-toluidine, $(m\text{-}CH_3C_6H_4NCH_2—)_x$, is formed from formaldehyde and *m*-toluidine. In the presence of alkali, bis(*o*-tolylamino)methane (b.p. 220–8°C. at 58 mm.) is obtained (7). In acid solution, with half a mole of formaldehyde, 4,4′-diamino-2,2′-dimethyldiphenylmethane (m.p. 123°C.) is obtained. Under more drastic conditions, resins result (141).

In neutral solution, under ordinary conditions, formaldehyde and *p*-toluidine form more or less indefinite products, among which is anhydroformaldehyde *p*-toluidine, $(p\text{-}CH_3C_6H_4NCH_2—)_x$ (m.p. 127–8°C.), in which x is probably 3 (6, 57, 79, 93, 152). In the presence of alkali, bis(*p*-tolylamino)methane is formed (m.p. 89°C.) (35, 37).

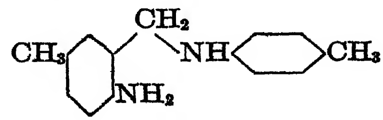
A considerable amount of attention has been devoted to the study of the reaction between *p*-toluidine and formaldehyde in acid solution, especially by Wagner and coworkers (42, 114, 143). The great interest in this particular problem first arose from the isolation by Troeger (121) of a base of the formula $C_{17}H_{18}N_2$, since known as "Troeger's base." The same base was obtained by Löb, during the electrolytic reduction of *p*-nitrotoluene in acid solution in the presence of formaldehyde (79), although the identity of the two substances was not realized until later (56). A number of other substances are produced at the same time as Troeger's base (42, 75, 84). Among these the following have definitely been identified: *N*-methyl-*p*-toluidine, *N*,*N*-dimethyl-*p*-toluidine, 1-(*p*-tolylaminomethyl)-2-hydroxy-3-*p*-tolyl-6-methyl-1,2,3,4-tetrahydroquinazoline, and 3-*p*-tolyl-6-methyl-3,4-dihydroquinazoline. The follow-

[2] *p*,*p*′-Diaminodiphenylmethane may easily be prepared by reacting aniline with half a mole of formaldehyde in weakly acid solution (142).

ing mechanism has been proposed to account for the products observed (42, 114, 149):

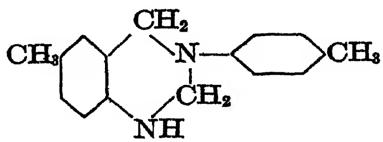
(*a*) $2CH_3C_6H_4NH_2 + CH_2O \rightarrow \underset{\text{I}}{CH_3C_6H_4NHCH_2NHC_6H_4CH_3} + H_2O$

(*b*) $\text{I} + CH_3C_6H_4NH_2 + CH_3C_6H_4NH_3Cl \xrightarrow{\text{rearrangement}}$



II

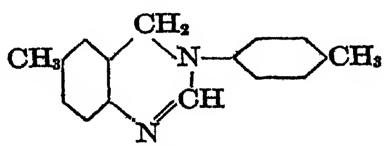
(*c*) $\text{II} + \text{I} \rightarrow$



$+ 2CH_3C_6H_4NH_2$

3-*p*-Tolyl-6-methyl-1,2,3,4-tetrahydroquinazoline
III

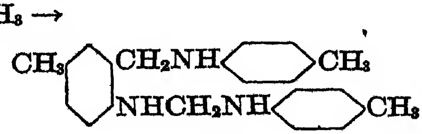
(*d*) $\text{III} + \text{I} \xrightarrow{H^+}$



3-*p*-Tolyl-6-methyl-3,4-dihydroquinazoline
IV

$+ CH_3C_6H_4NHCH_3 + CH_3C_6H_4NH_2$
N-methyl-*p*-toluidine *p*-Toluidine

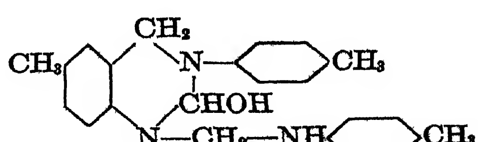
(*e*) $\text{II} + CH_2O + NH_2C_6H_4CH_3 \rightarrow$



$+ H_2O$

V

(*f*) $\text{V} + HCOOH \rightarrow$

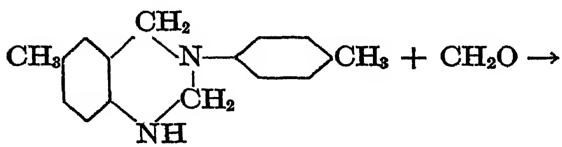


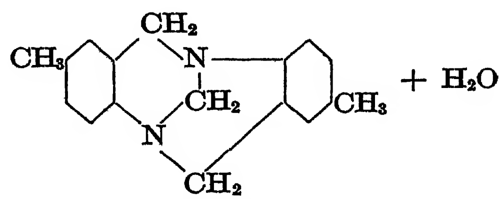
$+ H_2O$

1-(*p*-Tolylaminomethyl)-2-hydroxy-3-*p*-tolyl-6-methyl-1,2,3,4-tetrahydroquinazoline
VI

(The formic acid required for the last step may be formed during the reaction by oxidation of part of the formaldehyde present in the reaction mixture.)

The structure of Troeger's base has recently been established by Spielman (115), who has likewise accomplished its synthesis from 3-*p*-tolyl-6-methyl-1,2,3,4-tetrahydroquinazoline and formaldehyde in acid solution.





Troeger's base

The mechanism by which the base is formed from *p*-toluidine probably involves this step also (141).

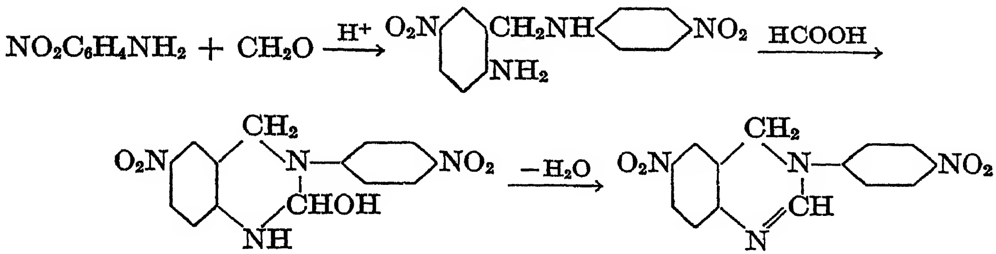
(*c*) *Nitroanilines.* The bis(nitrophenylamino)methanes are produced when the nitroanilines are boiled with formaldehyde in the presence of aqueous alcohol (101). The melting points of the *o*-, *m*-, and *p*-isomers are 195°, 213°, and 232°C., respectively.

When heated with formaldehyde and concentrated hydrochloric acid, *o*-nitroaniline is converted to a polymeric substance having the composition of an anhydro-*m*-nitro-*p*-aminobenzyl alcohol, $(C_7H_6N_2O_2)_x$. The same substance is obtained by treating bis(*o*-nitrophenylamino)methane with hydrochloric acid, a mole of *o*-nitroaniline being eliminated in the transformation (91).

$$x\mathrm{CH_2(NHC_6H_4NO_2)_2} \xrightarrow{\mathrm{HCl}} (\mathrm{C_7H_6N_2O_2})_x + x\mathrm{NH_2C_6H_4NO_2}$$

The reaction of *p*-nitroaniline with formaldehyde in the presence of acid is somewhat more complex. A bis(phenylamino)methane is not isolated in this case, but its semidine rearrangement product is: namely, 4-nitrophenylamino-2′-amino-5′-nitrophenylmethane (92). In the presence of formic acid, this condenses further to give 2-hydroxy-3-(*p*-nitrophenyl)-6-nitro-1,2,3,4-tetrahydroquinazoline, which, in turn, readily

undergoes dehydration to 3-(*p*-nitrophenyl)-6-nitro-3,4-dihydroquinazoline (84).



Under more severe conditions complex substances are formed, which correspond approximately to polymeric dehydration products of the crystalline intermediates mentioned above (92).

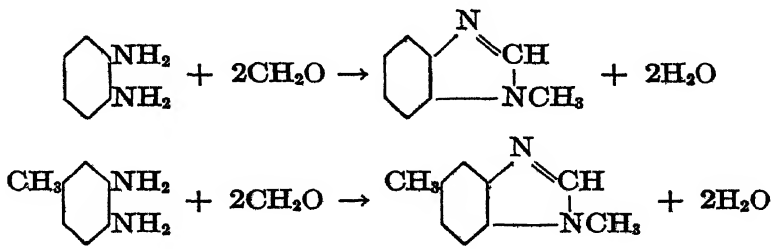
(*d*) *Anisidines and phenetidines.* The *o*- and *p*-anisidines condense with alcoholic formaldehyde in the presence of alkali to give the corresponding bis(methoxyphenylamino)methanes, m.p. 86° and 66°C. In the absence of alkali, however, *p*-anisidine gives the cyclic trimer of methylene-*p*-anisidine, m.p. 132°C. (7). In the case of *p*-phenetidine, bis(*p*-ethoxyphenylamino)methane, m.p. 75°C., is formed, either in the presence or in the absence of alkali (7, 144). In the presence of acid, the products include *N*-methyl-*p*-phenetidine and 3-(*p*-ethoxyphenyl)-6-ethoxy-3,4-dihydroquinazoline (84). The dihydroquinazolines may be obtained from *p*-anisidine and *p*-phenetidine by rearranging the diphenylaminomethanes to the corresponding semidine bases (by heating with the amine and its hydrochloride) and reacting this product with formic acid (144).

(*e*) *Aminobenzoic acids.* The isomeric (*o*-, *m*-, and *p*-) bis(carboxyphenylamino)methanes (m.p. 158°, 129°, and 167°C.) are obtained by treating the aminobenzoic acids with formaldehyde in alcohol (7). The formation of resins was also noted, especially in the case of *m*-aminobenzoic acid.

(*f*) *Chloro- and bromo-anilines.* The isomeric chloroanilines give the corresponding *o*-, *m*-, and *p*-bis(chlorophenylamino)methanes (m.p. 84°, 73°, and 65°C.) when treated with formaldehyde and alkali in alcohol (7). Without the alkali, the trimeric methylenechlorophenylimines were obtained from *m*- and *p*-chloroanilines. Bis(*p*-bromophenylamino)methane may be similarly prepared (144). In the presence of acid, however, the main products obtained from *p*-chloro- and *p*-bromo-anilines are the 3,4-dihydroquinazolines, m.p. 192° and 205.8°C., provided the temperature is kept below room temperature (145). At a higher temperature (70°C.) the main product obtained from *p*-bromoaniline was the ketone 3-(*p*-bromophenyl)-6-bromo-3,4-dihydro-4-quinazolone (12).

2. Formaldehyde and aromatic diamines

A somewhat unusual type of reaction occurs when formaldehyde is mixed with weakly acid solutions of *o*-phenylenediamine or *m*-methyl-*o*-phenylenediamine. The products are *N*-methylbenzoimidazoles, formed by condensation of two molecules of the aldehyde with one molecule of the diamine (46).



In neutral solution, however, products were obtained which corresponded empirically to dimers of the corresponding Schiff bases (or bis(methyleneimino)benzenes). 1,2-Diaminonaphthalene gave a similar product in neutral solution.

The formation of the "aldol" *p,p'*-bis(hydroxymethylamino)diphenyl, $HOCH_2NHC_6H_4C_6H_4NHCH_2OH$, on treatment of benzidine with formaldehyde in alcohol solution has been reported (71). Earlier, however, the production of the Schiff base, $CH_2{=}NC_6H_4C_6H_4N{=}CH_2$, under these conditions was claimed (108).

3. Formaldehyde and secondary aromatic amines

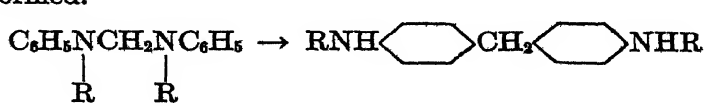
Secondary aromatic amines do not resinify in the presence of formaldehyde as readily as do the primary amines. Neither do they form the substituted hydroquinazolines. Consequently it is easier in these cases to isolate some of the simpler reaction products.

In neutral or slightly alkaline solution the alkyldiphenylaminomethanes are readily obtained, according to the following equation:

$$2RNHAr + CH_2O \rightarrow \underset{\displaystyle R}{\underset{|}{Ar{-}N}}{-}CH_2{-}\underset{\displaystyle R}{\underset{|}{N{-}Ar}} + H_2O$$

Examples are the bases derived from *N*-methylaniline, *N*-ethylaniline, *N*-methyl-*o*-toluidine, and *N*-methyl-*p*-toluidine (129).

When the reaction is carried out in the presence of hot acid, the benzidine type of rearrangement occurs and the *p*-alkylaminodiphenylmethanes are formed.



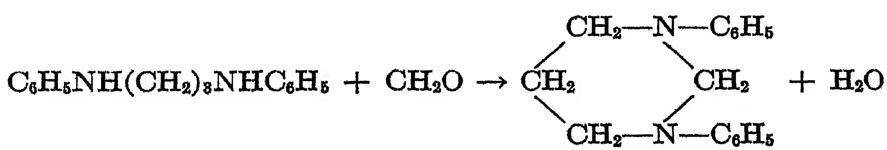
The latter are liquids, boiling in the neighborhood of 250–300°C. at about 10 mm. (55, 129, 141, 142).

The behavior of diphenylamine is very similar. In boiling benzene bis(diphenylamino)methane is obtained; on treatment with dilute acid, this undergoes the benzidine type of rearrangement to give bis(*p*-phenylaminodiphenyl)methane (18).

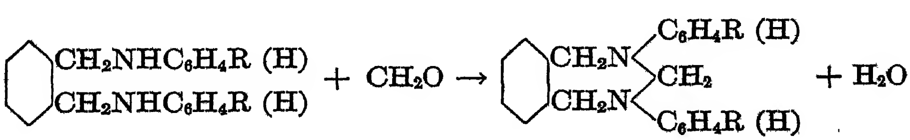
Under somewhat different conditions, the hydrochlorides of anhydro-*p*-alkylaminobenzyl alcohols can be isolated, and these can be converted to the free bases by suspending them in ice-cold alcohol and treating with 20 per cent sodium hydroxide (160). The anhydro-*p*-alkylaminobenzyl alcohols derived from methyl-, ethyl-, *n*-propyl-, *n*-butyl-, isoamyl-, and benzyl-anilines are solids of definite melting points. The molecular weights in benzene correspond almost exactly to those calculated for the trimers, but in camphor the apparent molecular weights are much higher (160). Molecular weights corresponding to those of dimers have also been reported in the cases of anhydromethyl-*p*-aminobenzyl alcohol and anhydroethyl-*p*-aminobenzyl alcohol (51).

In the case of arylated 1,2-diamines, the tetrahydroimidazoles are readily isolated, even when working with acid aqueous or alcoholic solutions. Examples are 1,3-diphenyltetrahydroimidazole (m.p. 124°C.), 1,3-di-(*p*-ethoxyphenyl)tetrahydroimidazole (m.p. 214°C.), 1,3-di-(*m*-tolyl)tetrahydroimidazole (m.p. 100°C.), and 1,3-di(*p*-tolyl)tetrahydroimidazole (m.p. 176°C.) (6, 111). In these cases, then, the results are exactly similar to those obtained with aliphatic 1,2-diamines.

By a similar reaction, 1,3-diphenylhexahydropyrimidine (m.p. 87°C.) is formed when 1,3-bis(phenylamino)propane and formaldehyde are allowed to react in the cold (110, 128).



1,3-Bis(*p*-tolylamino)propane and formaldehyde give the homologous hexahydropyrimidine (m.p. 60–3°C.). Diphenylxylylenediamine and di-*p*-tolylxylylenediamine are said to form the corresponding seven-membered cyclic compounds when heated with formaldehyde (111).



B. Acetaldehyde

1. Acetaldehyde and aromatic monoamines

This reaction, like that of aniline and formaldehyde, has been the subject of extensive study over a period of about sixty years. A number of crystalline products have been recognized, as well as the usual amorphous and resinous by-products.

If aniline and acetaldehyde are allowed to react in the cold, either in neutral solution or in the presence of sodium hydroxide, the product obtained is 1,1-bis(phenylamino)ethane, $C_6H_5NHCH(CH_3)NHC_6H_5$, m.p. 51°C. (36, 93). Under less carefully controlled conditions, amorphous and oily products are formed along with crystalline substances (106, 119). When the reaction is carried out in cold alcohol, in the presence of hydrochloric acid, two crystalline substances are found, both of which are dimers of the hypothetical "ethylideneaniline," $C_6H_5N{=}CHCH_3$ (132, 133). The higher melting of these isomers ("Eckstein's base," m.p. 126°C.) has been assigned the following structure:

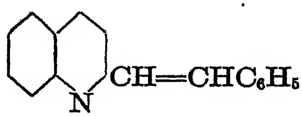
$$\begin{array}{l} C_6H_5NH\underset{\displaystyle CH_3}{\underset{|}{C}}H{-}CH{=}CH{-}NHC_6H_5 \end{array}$$

and the assumption is made that the isomeric base (m.p. 85°C.) is a *cis–trans* isomer (38). Both dimeric products may therefore be conceived of as formed by condensation of two molecules of aniline with one molecule of acetaldol, followed by a 1,3-shift of hydrogen (39).

$$\begin{array}{ccc}
CH_3CHOHCH_2CHO + H_2NC_6H_5 & \rightarrow & CH_3\underset{\displaystyle NHC_6H_5}{\underset{|}{C}}HCH_2CH{=}NC_6H_5 + 2H_2O \\
+ & & \downarrow \text{rearrangement} \\
H_2NC_6H_5 & & CH_3\underset{\displaystyle NHC_6H_5}{\underset{|}{C}}HCH{=}CHNHC_6H_5
\end{array}$$

Further complications are noted if aniline hydrochloride and acetaldehyde are allowed to stand in water solution for several weeks. Under these circumstances at least two more amorphous products are formed, both of which correspond empirically to higher polymers of ethylideneaniline (40, 130). These polymers (generally designated as "Schultze's base") are probably derived from Eckstein's base, since the latter can be converted to Schultze's base by treatment with dilute hydrochloric acid at room temperature. Eckstein's base also reacts with benzaldehyde when heated, but the product thus obtained is a definitely crystalline

material, and probably has the structure of a benzalquinaldine—namely, α-styrylquinoline (40).



Benzylaniline is also formed during this reaction.

2. Acetaldehyde and other aromatic amines

The diphenylaminoethanes are readily obtained from para-substituted anilines. Thus *p*-nitro- and *p*-chloro-anilines, with acetaldehyde in ether, give 1,1-bis(*p*-nitrophenylamino)ethane and 1,1-bis(*p*-chlorophenylamino) ethane (m.p. 167° and 65°C.) (37).

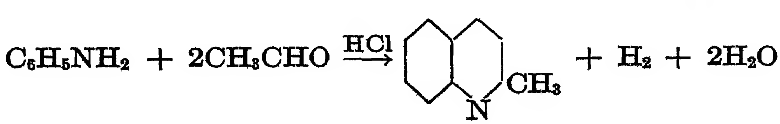
C. Homologs of formaldehyde

1. Homologs of formaldehyde and aromatic amines

Aniline and propionaldehyde form a dimeric product which is probably similar in structure to Eckstein's base (139). Valeraldehyde and aniline also give a dimeric product (134). Butryaldehyde, isobutyraldehyde, isovaleraldehyde, and heptaldehyde form monomeric anhydro products with aniline (presumably of the Schiff base type), as does isovaleraldehyde with *p*-toluidine. Isobutyraldehyde also gives some 1,1-bis(phenylamino)-isobutane (50, 77, 135, 137, 138). Isovaleraldehyde, with *p*-chloroaniline in ether, gives a dimer of the Schiff base, but with *p*-nitroaniline, under the same conditions, 1-bis(*p*-nitrophenylamino)-3-methylbutane, $(CH_3)_2CHCH_2CH(NHC_6H_4NO_2)_2$, is formed (39).

2. The formation of quinoline bases from aliphatic aldehydes and aromatic amines

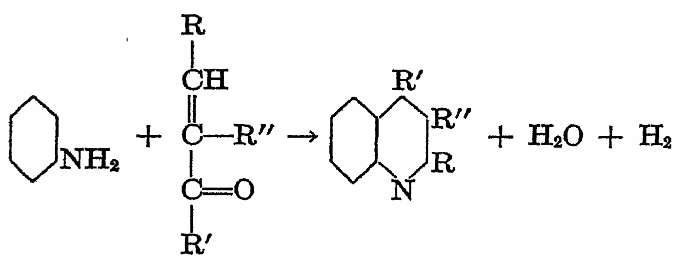
When aromatic amines are heated with aliphatic aldehydes and concentrated hydrochloric acid, a reaction occurs which bears considerable resemblance both to the Skraup quinoline synthesis and to the Tschitschibabin reaction. Ring closure is effected, oxidation and dehydration occur, and a quinoline derivative is formed. In the case of paraldehyde and aniline, the reaction may be represented by the following scheme (31):



The product, α-methylquinoline, or quinaldine, can also be formed from aniline and aldol. The toluidines react in a similar fashion, giving 2,8-, 2,7-, and 2,6-dimethylquinolines, respectively. Among other substituted anilines which react similarly are sulfanilic acid, *o*-aminophenol, *o*-anisidine, *p*-anisidine, cumidine, α-naphthylamine, and β-naphthylamine (the last two giving the corresponding benzoquinaldines) (32).

Higher aldehydes react with aniline to give the higher alkylated quinolines. Thus, propionaldehyde yields 2-ethyl-3-methylquinoline; butyraldehyde yields 2-propyl-3-ethylquinoline; isovaleraldehyde yields 2-butyl-3-propylquinoline; and heptaldehyde yields 2-hexyl-3-amylquinoline (33). Tiglic aldehyde, $CH_3CH{=}C(CH_3)CHO$, condenses with aniline to give 2,3-dimethylquinoline, only one molecule of water being eliminated in this case (104). The action of a mixture of propionaldehyde and formaldehyde on aniline leads to a mixture of 3-methylquinoline and 2-ethyl-3-methylquinoline (131). These products would be expected from the action of α-methylacrolein and of 2-methylpenten-2-al upon aniline. The inference is clearly that it is the "crotonaldehyde" of the aldehyde (or mixture of aldehydes) which actually condenses with the amine. The reaction may similarly be carried out with mixtures of aldehydes and ketones (5). A mixture of paraldehyde and acetone thus gives 2,4-dimethylquinoline when heated with aniline; formaldehyde and acetone give 2-methylquinoline; and acetaldehyde and acetophenone give 2-methyl-4-phenylquinoline. (Cinnamaldehyde gives 2-phenylquinoline, rather than 4-phenylquinoline. In this case, something akin to the α,γ-transformation met with in the case of the Claisen rearrangement of allyl phenyl ethers probably occurs.)

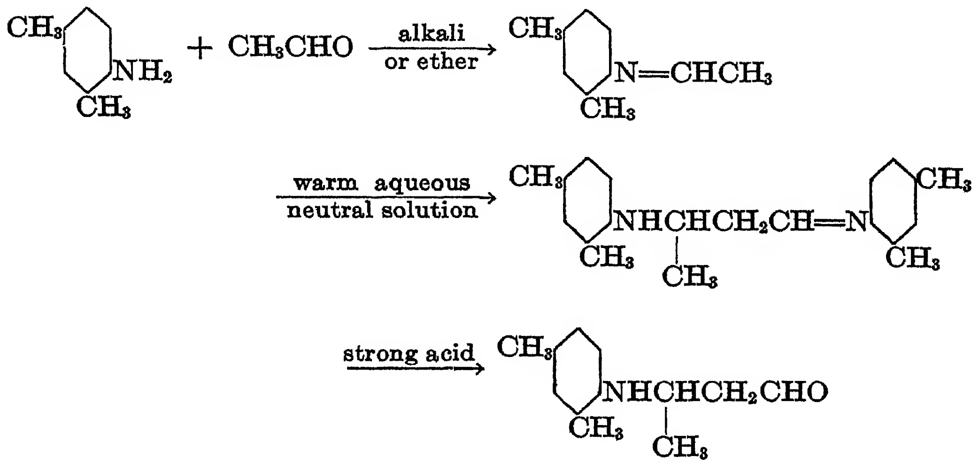
A general statement of the quinaldine reaction, as understood when Von Miller was carrying out his classical work on the subject, would be somewhat as follows:



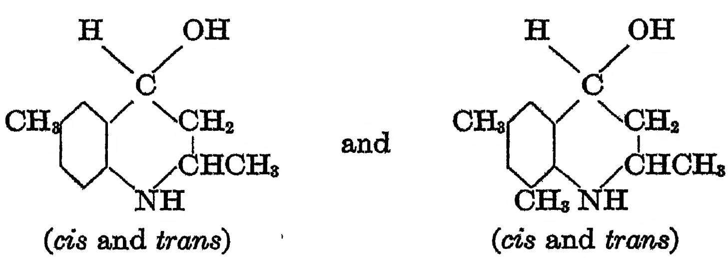
On this view, it should be possible to prepare 2,3,4-trialkyl (or alkyl aryl) derivatives of quinoline by this method. Apparently the validity of this conclusion has not yet been proved or disproved.

Panajotow (97a) found that when *as*-*m*-xylidine is heated with concen-

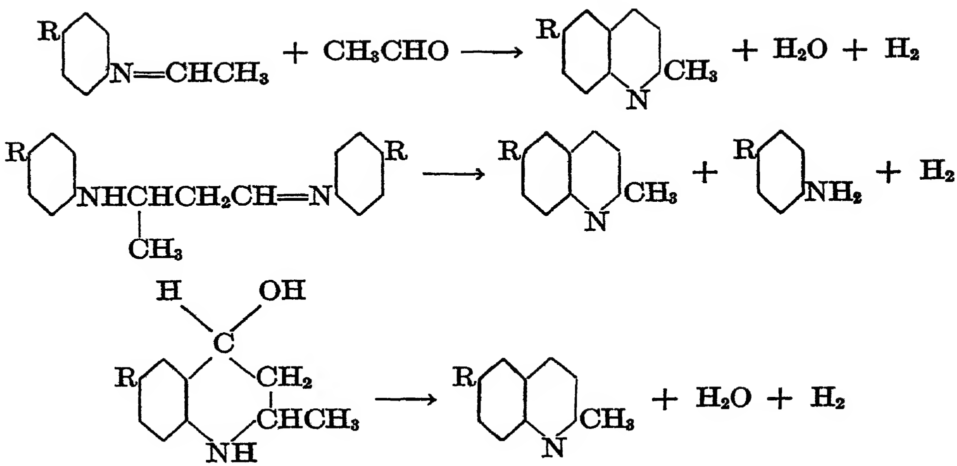
trated hydrochloric acid and paraldehyde the product is 2,6,8-trimethylquinoline, as would be expected on the basis of the above scheme. Von Miller and Plöchl (132a) reëxamined this reaction and found that at least three intermediates could be isolated. Thus:



The first is the simple Schiff base; the second is the Schiff base of the third, which, in turn, may be considered to be formed by condensation of one molecule of amine and one molecule of aldol. The intermediate formed in acid solution was found to exist in two stereoisomeric forms. These were later studied by Edwards, Garrod, and Jones (35a), who showed that neither had the above aldehyde structure, and that they were, in fact, *cis–trans* isomers of 2-methyl-4-hydroxy-1,2,3,4-tetrahydroquinoline. The compounds obtained in acid solution in the cases of *p*-toluidine and *as-m*-xylidine therefore have the following structures:

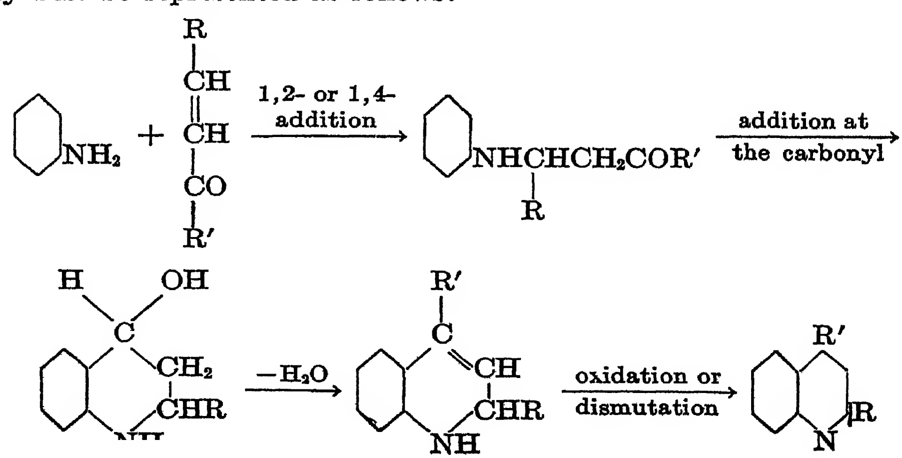


All of the intermediates isolated by Von Miller and Plöchl (132a) may be converted to quinoline derivatives by heating, alone or with acidic agents. The changes involved are represented schematically as follows:



Actually, however, very little, if any, free hydrogen can ever be detected. When the above changes were studied more closely, it was observed that the dihydroquinoline derivative which might be expected if no hydrogen is evolved is converted, by a dismutation reaction, into an equilibrium mixture of a quinoline and a tetrahydroquinoline derivative (65a). Under the conditions of the quinaldine synthesis *per se*, however, tetrahydroquinoline derivatives are not found (93a). Instead the dihydroquinoline is undoubtedly oxidized *in situ*, the oxidizing agents being the simple Schiff bases which are always present during a quinaldine synthesis. (In the case of aniline, for example, ethylideneaniline and crotylideneaniline are presumably present, and are reduced to ethylaniline and *n*-butylaniline by the dihydroquinoline, which is thereby oxidized to quinaldine (93a).)

On the basis of these facts, the mechanism of the quinaldine reaction may best be represented as follows:



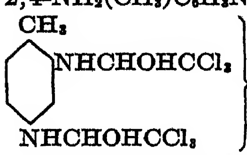
However, this scheme must still be considered as somewhat speculative, since no indisputable proof of it has yet been brought forward.

D. Chloroacetaldehydes and aromatic amines

The initial addition stage (the "aldol" stage) of the amine–aldehyde reaction may be arrested when α-chlorinated acetaldehydes are treated with aromatic amines. The presence of α-chlorine apparently stabilizes the addition compound with respect to both hydrolysis and further reaction with the reagents present. This is especially true in the case of chloral. The α-hydroxy-β-trichloroethylarylamines are therefore con-

TABLE 2

Addition compounds of chloral and aromatic amines

ADDITION COMPOUND	MELTING POINT	REFERENCE
	°C.	
p-$CH_3C_6H_4NHCHOHCCl_3$	75	(37)
p-$NO_2C_6H_4NHCHOHCCl_3$	128	(37)
2,4-$NO_2(CH_3)C_6H_3NHCHOHCCl_3$	187–188	(155)
3,4-$Cl(CH_3)C_6H_3NHCHOHCCl_3$	182–183	(155)
α-$C_{10}H_7NHCHOHCCl_3$	93–93.5	(105)
β-$C_{10}H_7NHCHOHCCl_3$	101	(105)
o-$NH_2C_6H_4NHCHOHCCl_3$	72	(105)
p-$NH_2C_6H_4NHCHOHCCl_3$	80	(105)
3,4-$NH_2(CH_3)C_6H_3NHCHOHCCl_3$	86	(105)
2,4-$NH_2(CH_3)C_6H_3NHCHOHCCl_3$	67–68	(105)
	56–57	(105)

The α-amino alcohol derived from aniline and chloral was obtained in the form of the double compound with chloral hydrate (m.p. 56.5°C.) (105).

veniently prepared by mixing the reactants in warm inert solvents, such as ether, chloroform, benzene, or toluene. The addition compound usually precipitates under these conditions.

Table 2 shows examples of addition compounds of chloral and aromatic amines.

In the majority of cases, however, the α-amino alcohols condense with a second molecule of amine, even under the mild conditions mentioned above, with the result that the bis(phenylamino)methane is obtained. These condensation products are almost all crystalline solids of fairly high melting points. Examples of this type of condensation product of chloral

and aromatic amines are given in table 3. The bis(phenylamino)methanes derived from dichloroacetaldehyde and aniline and from chloroacetaldehyde and *p*-chloroaniline have also been described (37). However, it is stated that 2,4,6-trichloroaniline and 2,6-dichloro-4-nitroaniline do not react with chloral (37).

TABLE 3

Condensation products of chloral and aromatic amines

CONDENSATION PRODUCTS	REFERENCE
Bis(phenylamino)trichloromethylmethane	(37, 109, 146)
Condensation products from chloral and the toluidines	(37, 147, 154)
Condensation products from chloral and the nitroanilines	(37, 156)
Condensation products from chloral and halogenated anilines	(37, 154, 155)
Condensation products from chloral and the anisidines	(154)
Condensation products from chloral and the aminobenzoic acids	(96, 154, 155)
Condensation products from chloral and the nitrotoluidines	(155)
Condensation products from chloral and nitrobromoanilines	(155)

V. REACTION OF AROMATIC ALDEHYDES AND AROMATIC AMINES

A. Aromatic aldehydes and primary aromatic amines

1. Schiff bases

Under ordinary conditions aromatic aldehydes and aromatic amines react very readily to give Schiff bases. These Schiff bases are well-defined, crystalline substances. In general they do not react further with either of the reagents used in their preparation, as do most of the other types of simple intermediates considered thus far. They may accordingly be looked upon as stable representatives of the "crotonaldehyde stage" of the amine–aldehyde condensation.

The aromatic Schiff bases are obtained merely by warming the aldehyde and amine together. However, it is sometimes more convenient to work in a solvent such as alcohol, dilute acetic acid, or glacial acetic acid. Sometimes the reaction is aided by a trace of acid; in other cases the hydrochlorides of the amines can be used in the synthesis.

So many aromatic Schiff bases are described in the literature that it is virtually impossible to list even a representative number of them. Table 4 gives a partial insight into the very great amount of work which has been done in this field.

2. Velocity of formation of Schiff bases

A study was made of the comparative rates of formation of Schiff bases from aniline and substituted anilines and aromatic aldehydes by Oddo and

Tognacchini, using a cryoscopic method to follow the course of the reaction (97). It was noted that for the nitroanilines the speeds of formation were in the order $o > m > p$. *p*-Tolualdehyde reacted much faster than phenylacetaldehyde, and cinnamaldehyde was much more rapid than either. Cinnamaldehyde was also much more rapid than *m*-nitrocinnamaldehyde. The speed of reaction decreased in the series: cinnamaldehyde, anisaldehyde, vanillin, piperonal. The results were not amenable to a general theoretical interpretation.

TABLE 4

Schiff bases formed from aromatic aldehydes and aromatic amines

SCHIFF BASES FORMED FROM		REFERENCE
Aldehyde	Amine	
Benzaldehyde	Aniline	(8, 15)
Benzaldehyde	Substituted anilines	(13, 14a, 29, 58, 60, 76, 94, 148, 149)
Chlorobenzaldehydes	Aniline or *p*-chloroaniline	(60)
o-Nitrobenzaldehyde	Substituted anilines	(14a, 76, 94, 148)
m-Nitrobenzaldehyde	Substituted anilines	(14a, 60, 94, 113, 148)
p-Nitrobenzaldehyde	Substituted anilines	(14a, 43, 60, 67, 82, 94, 113, 148)
2,4-Dinitrobenzaldehyde	Substituted anilines	(81, 83)
2,4,6-Trinitrobenzaldehyde	Substituted anilines	(80)
2,4,6-Trimethylbenzaldehyde	Aniline; *p*-chloroaniline	(60)
Salicylaldehyde	Substituted anilines	(58, 60, 76, 148)
p-Hydroxybenzaldehyde	Aniline	(29)
o-Methoxybenzaldehyde, *p*-Methoxybenzaldehyde	*p*-Aminophenol	(58)
p-Aminobenzaldehyde	Substituted anilines	(94, 150)
p-Dimethylaminobenzaldehyde	Substituted anilines	(94, 151)
p-Diethylaminobenzaldehyde	*p*-Dimethylaminoaniline	(94)
Furfural	Aminophenols	(76)
Terephthalaldehyde	Aminophenols	(76)
Cinnamaldehyde	Aminophenols	(58, 76)
Vanillin	Substituted anilines	(14a, 76)
Piperonal	2-Methyl-5-aminobenzonitrile	(14a)

3. The addition stage

The products of the "aldol" stage of the reaction have been arrested in relatively few instances when both reactants are aromatic. Those isolated include the α-amino alcohol from *m*-aminobenzoic acid and benzaldehyde, *m*-$HOOCC_6H_4NHCHOHC_6H_5$ (59); that from *p*-nitroaniline and benzaldehyde, *p*-$NO_2C_6H_4NHCHOHC_6H_5$ (140); and that from 2,4,6-trinitrobenzaldehyde and aniline, $(NO_2)_3C_6H_2CHOHNHC_6H_5$ (80).

Other aryl-substituted α-amino alcohols have been prepared indirectly by either of two methods. The Schiff bases, on treatment with dry hydrogen chloride in ether or benzene, can be converted to hydrochlorides, which can then be hydrolyzed to the α-amino alcohols by treatment with sodium carbonate solution at low temperatures (60). The α-amino alcohols derived from benzaldehyde and *p*-chloroaniline and *p*-bromoaniline were obtained in this manner.

$$C_6H_5CH{=}NC_6H_4X \xrightarrow{HCl}$$

$$\underset{HCl}{C_6H_5CH{=}\dot{N}C_6H_4X} \xrightarrow[Na_2CO_3]{H_2O} C_6H_5CHOHNHC_6H_4X$$

These hydrochlorides may be derivatives of chloromethylamine, i.e., ArCHClNHAr (60), but neither this structure nor the ammonium salt structure pictured above has been established with certainty.

The alternate method for preparing these substances is first to prepare the hydrochloride of the α-amino alcohol by treating the aldehyde with the hydrochloride of the amine in the presence of a trace of mineral acid, and subsequently to treat this hydrochloride with cold sodium carbonate (29).

$$C_6H_5CHO + ClNH_3C_6H_4X \xrightarrow{H^+}$$

$$\underset{HCl}{C_6H_5CHOH\dot{N}HC_6H_4X} \xrightarrow[Na_2CO_3]{H_2O} C_6H_5CHOHNHC_6H_4X$$

The addition products prepared in this way include those from benzaldehyde and *p*-nitroaniline, from salicylaldehyde and aniline, and from *p*-hydroxybenzaldehyde and aniline.

The α-amino alcohols are relatively unstable when both constituents are aromatic in nature. They lose water readily to give the Schiff bases. They cannot be kept for any length of time, for they are readily decomposed into their constituents on standing at room temperature. These two modes of decomposition are probably equilibrium reactions, but they have not yet been studied from that point of view.

B. Aromatic aldehydes and secondary aromatic amines

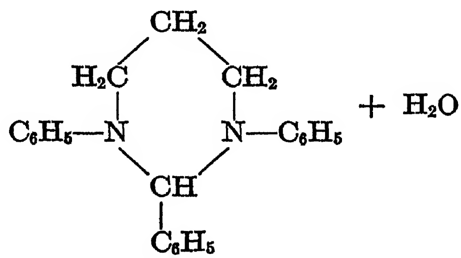
Schiff, in 1864, described the reaction product of *N*-ethylaniline and benzaldehyde (106) to which he assigned the structure,

$$\begin{array}{ccccccc} & & & C_6H_5 & & & \\ & & & | & & & \\ C_6H_5— & N & — & CH & — & N & —C_6H_5 \\ & | & & & & | & \\ & C_2H_5 & & & & C_2H_5 & \end{array}$$

and described it as a "resin" which showed no definite melting point. A description of a similar product, derived from methylaniline and benzaldehyde, cannot be found in the usual organic reference manuals. Very few, if any, crystalline derivatives of structures similar to that given above have been described.

Recently, however, it has been shown (128) that disecondary aromatic-aliphatic amines derived from propane can readily be converted to 1,3-diarylhexahydropyrimidines by treatment with the appropriate aromatic aldehyde in aqueous or alcoholic solution. As an example, 1,2,3-triphenylhexahydropyrimidine is thus formed from 1,3-bis(phenylamino)-propane and benzaldehyde.

$$C_6H_5NHCH_2CH_2CH_2NHC_6H_5 + C_6H_5CHO \longrightarrow$$

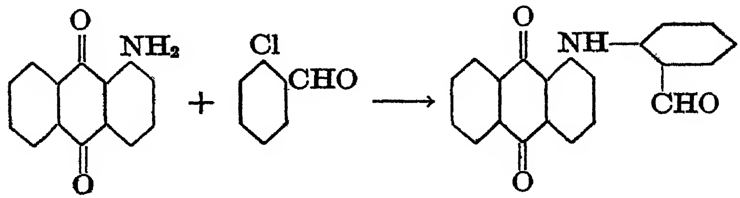


Other aldehydes which react readily are *p*-chlorobenzaldehyde, *p*-nitrobenzaldehyde, and furfural. However, salicylaldehyde, anisaldehyde, *p*-hydroxybenzaldehyde, and 5-methylfurfural failed to react. 1,3-Bis-(*p*-tolylamino)propane reacted readily with *p*-nitrobenzaldehyde, but failed to react with benzaldehyde.

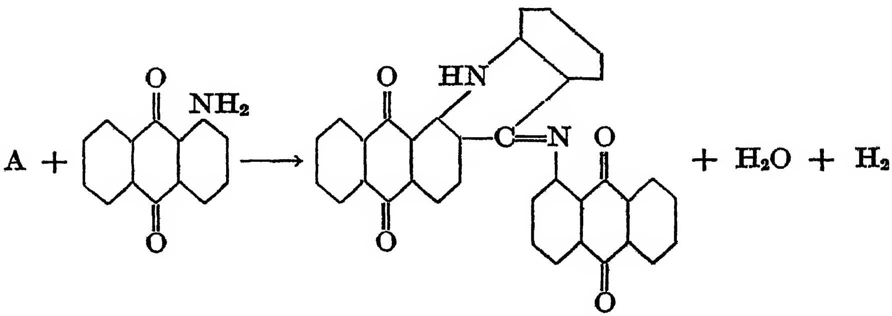
C. o-Chlorobenzaldehyde and aromatic amines

o-Chlorobenzaldehyde forms numerous Schiff bases, under the conditions outlined above. However, if the amine is heated with an excess of *o*-chlorobenzaldehyde in nitrobenzene or naphthalene, with the addition of a little "Naturkupfer C." and in the presence of sodium carbonate, an anomalous condensation occurs which leads to the formation of acridine derivatives (66, 88, 89, 90). The anomalous condensation is more easily realized in the case of very weakly basic amines, such as the aminoanthraquinones. The first phase of this reaction is a condensation of the Ullmann type, which leaves the aldehyde group intact. The second phase involves another molecule of amine, and effects ring-closure and con-

comitant oxidation. This may be illustrated, using 1-aminoanthraquinone as an example.

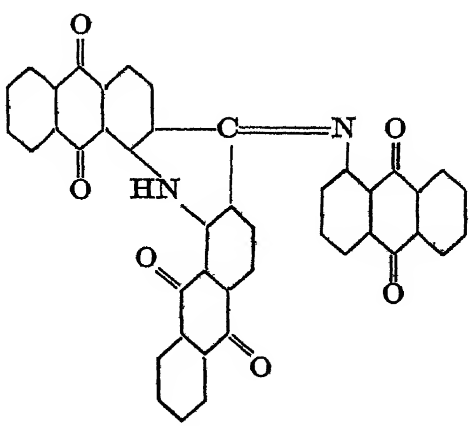


A

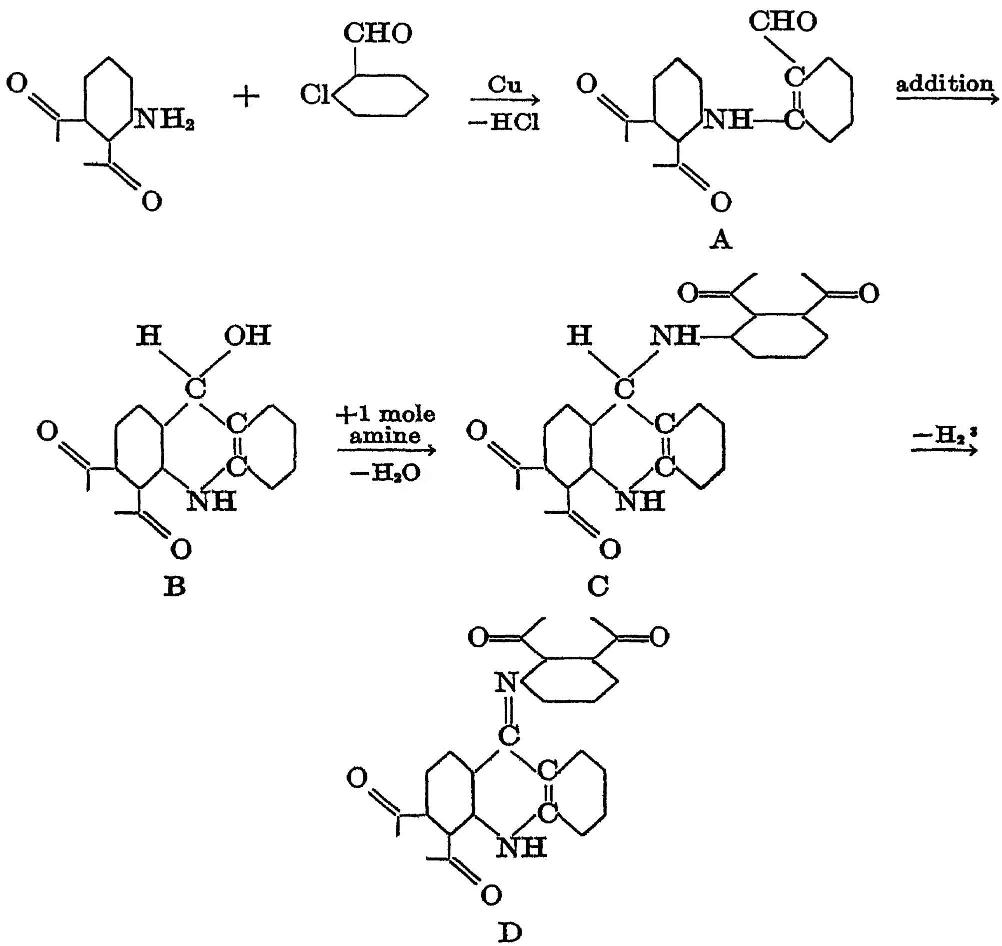


D

The product (D) is therefore a Schiff base derived from aminoanthraquinone and 3,4-phthalylacridone. 2-Aminoanthraquinone and 1-amino-5-chloroanthraquinone react similarly. 1-Aminoanthraquinone reacts with 1-chloroanthraquinone-2-aldehyde under the same conditions to give the Schiff base of 3,4,5,6-diphthalylacridone.



It is perhaps worthy of note that the intermediate A is similar in structure to a dehydrogenation product of the postulated intermediate in the quinaldine synthesis (page 322), and that the product D is, in one sense, a dihydroquinoline derivative. In fact, it has apparently been overlooked previously that this "anomalous" condensation bears a striking external resemblance to the quinaldine reaction. This becomes more apparent if postulated steps in the acridone synthesis are outlined as follows:

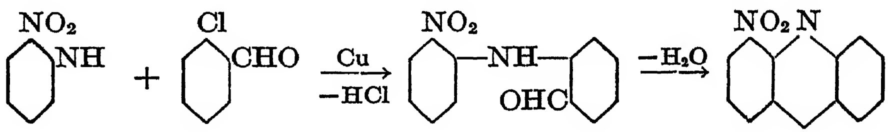


The intermediate A and the postulated intermediate B are structurally related to substances which would result in the quinaldine synthesis, if an acetylene aldehyde were used in place of the crotonaldehyde derivative

[3] Whether or not hydrogen gas is actually evolved is open to some question.

which is normally employed. The intermediate B, however, cannot lose water as readily as the intermediate formed in the quinaldine reaction and consequently it remains available for condensation with a second molecule of amine, giving C. Oxidation (or dehydrogenation) then gives the dihydroquinoline derivative D. The dihydroquinoline structure is therefore favored by the presence in the intermediate B of the condensed aromatic ring, and is fixed by conversion to the stable Schiff base. This mechanism is, of course, purely speculative, since the work necessary to prove or disprove its adequacy has not been undertaken.

Under the conditions of the above synthesis, the nitramines react in a somewhat more straightforward fashion with *o*-chlorobenzaldehyde. Simple acridines are formed in the second phase of the reaction by what is essentially a dehydration process. This may be illustrated by the case of *o*-nitroaniline, which gives 4-nitroacridine under these conditions (90).



Similarly, 2,4-dinitroaniline gives 2,4-dinitroacridine, 3-nitro-4-toluidine gives 3-nitro-2-methylacridine, and 2-nitro-4-chloroaniline gives 4-nitro-2-chloroacridine.

In many cases, however, the reaction fails to give either acridines or derivatives of acridone. The ordinary Schiff base is then usually obtained. This was observed in the cases of *o*-chlorobenzaldehyde and the following amines: *m*-nitroaniline, *p*-nitroaniline, 4-nitro-1-naphthylamine, 1-nitro-2-naphthylamine, *o*-, *m*-, and *p*-chloroanilines, 2,4-dichloroaniline, 2-chloro-4-toluidine, 2-aminofluorene, 2-aminofluorenone, 2-amino-7-nitrofluorene, 5-nitro-1-naphthylamine, 8-nitro-1-naphthylamine, 5-nitro-2-naphthylamine, and *o*-aminophenol. A similar result was observed in the case of 4-chloroaniline and 2-chloro-4-nitrobenzaldehyde. The reaction was observed to fail altogether in the cases of 2,4,6-trichloroaniline, 2,4-dinitro-1-naphthylamine, and 2,6-dinitroaniline (88, 89, 90).

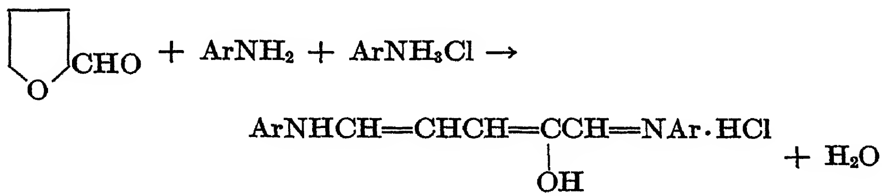
D. Furfural and aromatic amines

Stenhouse first discovered that furfural reacted with aniline in the presence of aniline hydrochloride and with toluidine in the presence of toluidine hydrochloride to give violet-colored salts (116). These reactions may be represented empirically as follows:

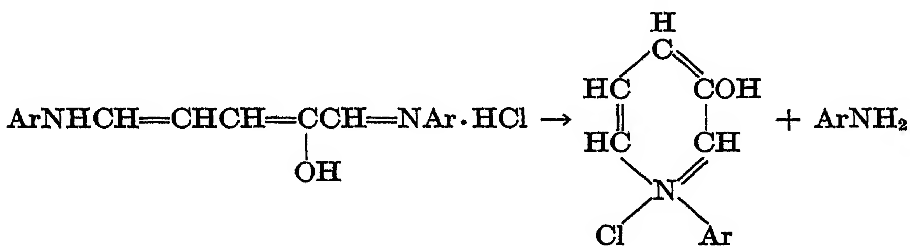
$$C_4H_3O\text{—}CHO + C_6H_5NH_2 + C_6H_5NH_3Cl \rightarrow C_{17}H_{18}N_2O_2 \cdot HCl$$

$$C_4H_3O\text{—}CHO + C_7H_7NH_2 + C_7H_7NH_3Cl \rightarrow C_{19}H_{22}N_2O_2 \cdot HCl$$

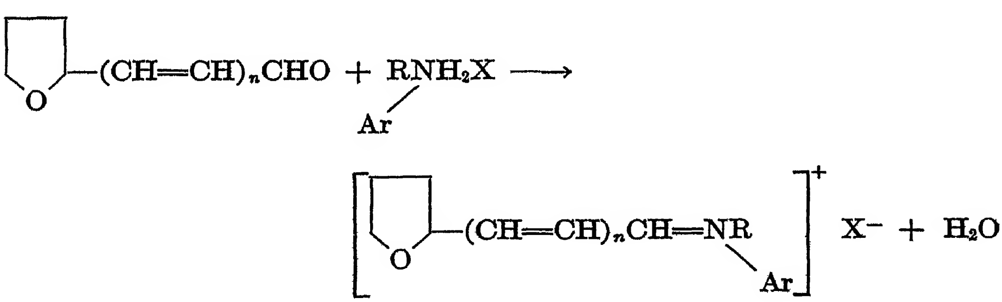
Subsequent investigations (28, 162) proved that these reactions involve scission of the furan ring by addition to it of 1 mole of the aryl amine; and, specifically, that the compound $C_{19}H_{22}N_2O_2 \cdot HCl$ is a hydrated salt of the Schiff base of an open-chain aldehyde, namely, 2-oxy-5-tolylamino-2,4-pentadienaltolylimine hydrochloride. The reaction scheme was therefore amplified as follows:

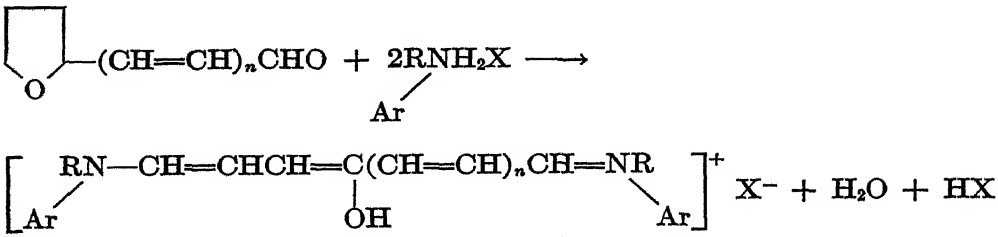


It was found, further, that when the hydrochloride was boiled with alcohol or acetic acid, a ring-closure occurred, leading, however, not to a furfural derivative again, but to a hydroxypyridinium aryl halide.



An elaboration of this investigation demonstrated that furfural could react as above with 2 moles of an arylamine to give an open-chain dye salt; or could react without ring opening with 1 mole of the aryl amine, in which case the product was also a deeply colored salt (72, 73). Similar dyestuffs could be obtained from vinyl homologs of furfural. From these facts the general reaction scheme may be formulated as follows:





(X^- may be any suitable anion, such as halogen or perchlorate)

[In those cases in which R is hydrogen, the 1:1 (closed ring) product is obviously a salt of the Schiff base derived from the amine and furfural (or vinyl homolog). In such cases the salt may be prepared from the Schiff base, merely by suspending the latter in alcohol and treating it with perchloric or other mineral acid. This procedure is, of course, in accord with the general method of preparing salts of Schiff bases previously discussed (60).]

This type of reaction is of quite general applicability. It has been applied to furfural, α-streptomonovinylenefurfural, α-streptodivinylenefurfural, and α-streptotrivinylenefurfural. Among the amines which react readily are *p*-anisidine, methyl-*p*-anisidine, tetrahydroquinoline, dihydro-α-methylindole, methylaniline, tetrahydro-*p*-toluquinoline, thalline, and aniline. It has recently been shown that furan ketones react according to the scheme outlined above (10).

VI. AMINE–ALDEHYDE RESINS

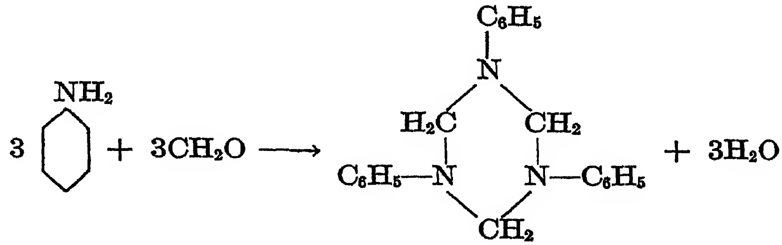
With the exception of the aromatic Schiff bases, which are quite stable substances, almost every other type of aldehyde–amine reaction product is capable of transformation, under the proper conditions, to amorphous or resinous materials of more or less indeterminate structure. The occurrence of such products is, in fact, one of the problems which has continually perplexed workers in this field. The chemistry of the amine–aldehyde resins has been very imperfectly worked out. Relatively brief attention will therefore be paid to this phase of the problem.

A. *Resins obtained from aniline and formaldehyde*

This particular subject has recently been discussed by Frey (49).

In neutral or slightly acid solution, the reaction of formaldehyde upon aniline probably goes through the amino alcohol (aldol) and imine (Schiff

base) stages. The product usually isolated under these conditions, as has been shown previously, is anhydroformaldehyde aniline trimer.



Simultaneously, polymer homologs of anhydroformaldehyde aniline are frequently observed. These probably result from chain-like polymerization of the monomeric Schiff base, much in the manner that polystyrene is formed from styrene monomer.

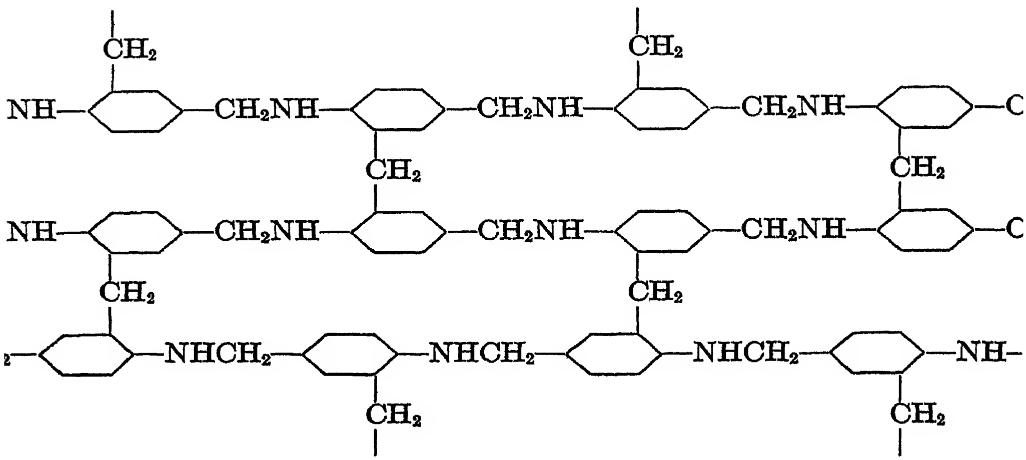
$$n[C_6H_5N{=}CH_2] \longrightarrow \text{-}\,\text{-}\underset{\displaystyle C_6H_5}{\underset{|}{N}}\text{—}CH_2\text{—}\left[\text{—}\underset{\displaystyle C_6H_5}{\underset{|}{N}}\text{—}CH_2\text{—}\right]_{n-1}\text{-------}$$

The nature of the groups on the end of the chains has not been determined.

When aniline reacts with an equimolecular proportion of formaldehyde in aqueous solution in the presence of about 1 mole of strong acid, or when anhydroformaldehyde aniline is treated with the equivalent quantity of strong acid, a resin of quite different character is formed. Since this resin may also be obtained by dehydrating aminobenzyl alcohol, it has been referred to as "anhydroaminobenzyl alcohol." By means of reduction studies, it has been proved that the methylene groups are attached to the nitrogen in anhydroformaldehyde aniline, but to the nucleus in anhydroaminobenzyl alcohol. It seems probable that the latter is also a linear polymer and that its structure may be represented as follows:

The nature of the end groups is again an open question, nor is the chain length known with certainty. It is likely that any product actually obtained consists of a mixture of polymer homologs, such as is known to be true in the case of the acid-catalyzed phenol–formaldehyde resins ("Novolaks"). In fact, products of relatively low chain length (n around 5 to 10) are readily obtained by using an excess of aniline, just as in the case of the analogous phenol–formaldehyde resins the use of an excess of phenol leads to products of low chain length. Again, in analogy with the same better known resins, such aniline–formaldehyde resins are permanently fusible and are easily soluble in many organic solvents.

When, however, an excess of formaldehyde is allowed to react upon aniline in acid solution, insoluble infusible products are formed. Hard, resistant materials of exceptional mechanical and electrical properties may thereby be obtained, by application of ordinary technical molding procedure. The structure of the "converted" or "cured" products is unknown, but it is generally supposed that these resins consist of long straight chains which are cross-linked at various points, that is, joined together by methylene groups at random positions in adjacent chains. Frey gives the following as an ideal pattern of this type of "cured" resin. Its similarity to the usual picture for cured or "C stage" Bakelite is worthy of note.



B. *Other amine–aldehyde resins*

Various substituted anilines have been caused to react with formaldehyde to form resins. The status of this particular subject, to 1935, is adequately treated in *The Chemistry of Synthetic Resins*, by Carleton Ellis, pages 686–711. To mention but a few typical examples: *o*-toluidine, *p*-toluidine, α-naphthylamine, β-naphthylamine, and *p*-chloroaniline all give brittle resins with formaldehyde. Monomethylaniline and dimethylaniline give fusible, brittle solids which have been used as copal substitutes. Aniline, the toluidines, naphthylamines, and xylidines have been condensed with acetaldehyde, aldol, crotonaldehyde, etc., to give liquids which may be used in varnishes. Aniline and furfural give dark-colored, brittle resins whose properties vary considerably with the method of preparation, but which have found some use in the lacquer and varnish industry.

A survey of more recent patent literature would show comparatively little important or concerted effort during the last few years. Only a few

patents need be cited to illustrate recent trends (163, 164, 165, 166, 167). The first of these provides an example of the technique used at present in the production of aniline–formaldehyde resins. The others describe (*a*) resinous reaction products obtained from aliphatic aldehydes and primary or secondary alkyl or aralkyl amines or secondary heterocyclic bases, which may be used for producing films, lacquers, or shellac substitutes; (*b*) condensation products of aliphatic aldehydes and polyalkylenepolyamines (such as diethylenetriamine and triethylenetetramine) which may be used as waterproofing agents; and (*c*) the reaction products of aliphatic aldehydes and polyaminodiazines, such as 2,4,6-triaminopyrimidine, 2,4-diaminoquinazoline, 2,4,5,6-tetraaminopyrimidine, and 2,4-diamino-5,6-benzoquinazoline. These last-named products are said to combine the excellent chemical and electrical resistance of the aniline–formaldehyde resins with the equally desirable light color and resistance to light of the urea–formaldehyde resins.

Finally, it should be noted that almost any of the amine–aldehyde reactions leads to the formation of tarry or resinous products under certain conditions. Naturally, the properties and structures of these products vary over wide latitudes, depending upon the nature of the reactants, the experimental conditions, and the treatment to which these materials are subjected before they are submitted to evaluating tests.

REFERENCES

(1) Andree: Ber. **35**, 424 (1902).
(2) Aschan: Ber. **48**, 874 (1915).
(3) Béhal and Choay: Ann. chim. [6]**26**, 11 (1892).
(4) Bergmann and Miekeley: Ber. **57**, 662 (1924).
(5) Beyer: J. prakt. Chem. [2]**33**, 393 (1886).
(6) Bischoff: Ber. **31**, 3248 (1898).
(7) Bischoff and Reinfeld: Ber. **36**, 35 (1903).
(8) Borodin: J. prakt. Chem. **75**, 19 (1859).
(9) Borodin: Ber. **6**, 1253 (1873).
(10) Borsche, Leditschke, and Lange: Ber. **71**, 957 (1938).
(11) Butlerow: Ann. **115**, 322 (1860).
(12) Cairncross and Bogert: Collection Czechoslov. Chem. Commun. **7**, 58 (1936).
(13) Calm: Ber. **17**, 2938 (1884).
(14) Cambrier and Brochet: Bull. soc. chim. [3]**13**, 404 (1895).
(14a) Cândea and Macovski: Bull. soc. chim. [5]**5**, 1487 (1938).
(15) Cech: Ber. **11**, 248 (1878).
(16) Chancel: Bull. soc. chim. [3]**11**, 933 (1894).
(17) Chemische Fabrik v. Heyden: German patents 216,072 and 216,073.
(18) Craig: J. Am. Chem. Soc. **55**, 3723 (1933).
(19) Delépine: Bull. soc. chim. [3]**13**, 128 (1895).
(20) Delépine: Compt. rend. **125**, 951 (1897).
(21) Delépine: Bull. soc. chim. [3]**19**, 171 (1898).

(22) DELÉPINE: Compt. rend. **125,** 178 (1897); **126,** 343 (1898); **126,** 648 (1898).
(23) DELÉPINE: Compt. rend. **128,** 105 (1899).
(24) DELÉPINE: Compt. rend. **144,** 853 (1907).
(25) DELÉPINE AND RIVALS: Compt. rend. **129,** 520 (1899).
(26) DENNSTEDT AND ZIMMERMANN: Ber. **21,** 1553 (1888).
(27) DICKENSON AND RAYMOND: J. Am. Chem. Soc. **45,** 22 (1923).
(28) DIECKMANN AND BECK: Ber. **38,** 4122 (1905).
(29) DIMROTH AND ZOEPPRITZ: Ber. **35,** 984 (1902).
(30) DOBLER: Z. physik. Chem. **101,** 1 (1922).
(31) DOEBNER AND VON MILLER: Ber. **16,** 2464 (1883).
(32) DOEBNER AND VON MILLER: Ber. **17,** 1698 (1884).
(33) DOEBNER AND VON MILLER: Ber. **17,** 1712 (1884).
(34) DUDEN AND SCHARF: Ann. **288,** 218 (1895).
(35) EBERHARDT AND WELTER: Ber. **27,** 1804 (1894).
(35a) EDWARDS, GARROD, AND JONES: J. Chem. Soc. **101,** 1376 (1912).
(36) EIBNER: Ber. **30,** 1444 (1897).
(37) EIBNER: Ann. **302,** 334 (1898).
(38) EIBNER: Ann. **318,** 58 (1901).
(39) EIBNER: Ann. **328,** 121 (1903).
(40) EIBNER AND KOCH: Z. angew. Chem. **39,** 1514 (1926).
(41) EINHORN AND PRETTNER: Ann. **334,** 217 (1904).
(42) EISNER AND WAGNER: J. Am. Chem. Soc. **56,** 1938 (1934).
(43) FISCHER: Ber. **14,** 248 (1881).
(44) FISCHER: Ber. **26,** 467 (1893).
(45) FISCHER AND PRAUSE: J. prakt. Chem. [2]**77,** 129 (1908).
(46) FISCHER AND WRESZINSKE: Ber. **25,** 2711 (1892).
(47) FRANCHIMONT AND VAN EPPS: Rec. trav. chim. **15,** 169 (1896).
(48) FRANCIS: Ber. **42,** 2216 (1909).
(49) FREY: Helv. Chim. Acta **18,** 491–513 (1935).
(50) FRIEDJUNG AND MOSLER: Monatsh. **22,** 460 (1901).
(51) FRIEDLÄNDER: Monatsh. **23,** 973 (1902).
(52) FRITSCH: Ann. **286,** 6 (1895).
(53) GABRIEL AND LEUPOLD: Ber. **31,** 2832 (1898).
(54) GNEHM AND BÄNZIGER: Ber. **29,** 876 (1886).
(55) GNEHM AND BLUMER: Ann. **304,** 115 (1899).
(56) GOECKE: Z. Elektrochem. **9,** 470 (1903).
(57) GRÜNHAGEN: Ann. **256,** 288 (1889).
(58) HAEGELE: Ber. **25,** 2753 (1892).
(59) HANTZCH AND KRAFT: Ber. **24,** 3521 (1891).
(60) HANTZCH AND SCHWAB: Ber. **34,** 822 (1901).
(61) HASKELBERG: Chemistry & Industry **54,** 261 (1935).
(62) HELLER: Ber. **27,** 3097 (1894).
(63) HENRY: Bull. acad. roy. belg. [3]**26,** 206 (1893); **28,** 359 (1894); **29,** 26 (1895).
(64) HENRY: Bull. acad. roy. Belg. [3]**29,** 355 (1895).
(65) HENRY: Compt. rend. **120,** 839 (1895).
(65a) JONES AND EVANS: J. Chem. Soc. **99,** 334 (1911).
(66) KALISCHER AND MAYER: Ber. **49,** 1994 (1916).
(67) KING AND LOWY: J. Am. Chem. Soc. **46,** 757 (1924).
(68) KNUDSEN: Ber. **47,** 2700 (1914).
(69) KOLDA: Monatsh. **19,** 610 (1898).
(70) KOLOTOV: Bull. soc. chim. [2]**43,** 112 (1885).

(71) KONDO AND ISHIDA: J. Pharm. Soc. Japan **489**, 979 (1922); Chem. Abstracts **17**, 1456 (1923).
(72) KÖNIG: J. prakt. Chem. [2]**88**, 193 (1913).
(73) KÖNIG, HEY, SCHULZE, SILBERKNEIT, AND TRAUTMANN: Ber. **67**, 1274 (1934).
(74) KRONEBERG: J. Russ. Phys.-Chem. Soc. **48**, 305 (1916); Chem. Abstracts **11**, 582 (1917).
(75) LEPETIT, MAFFEI, AND MAIMERI: Gazz. chim. ital. **57**, 862 (1927).
(76) LEVI: Gazz. chim. ital. **59**, 544 (1929).
(77) LIPPMANN AND STRECKER: Ber. **12**, 831 (1879).
(78) LOB: Rec. trav. chim. **55**, 859 (1936).
(79) LÖB: Z. Elektrochem. **4**, 428 (1898).
(80) LOWY AND BALY: J. Am. Chem. Soc. **43**, 341 (1921).
(81) LOWY AND DOWNEY: J. Am. Chem. Soc. **43**, 346 (1921).
(82) LOWY AND KING: J. Am. Chem. Soc. **43**, 625 (1921).
(83) LOWY AND WESCOTT: J. Am. Chem. Soc. **42**, 849 (1920).
(84) MAFFEI: Gazz. chim. ital. **58**, 261 (1928).
(85) MANNICH AND DAVIDSEN: Ber. **69**, 2106 (1936).
(86) MANNICH, HANDKE, AND ROTH: Ber. **69**, 2112 (1936).
(87) MASON: Ber. **20**, 270 (1887).
(88) MAYER AND BAUSA: Ber. **54**, 16 (1921).
(89) MAYER AND LEVIS: Ber. **52**, 1641 (1919).
(90) MAYER AND STEIN: Ber. **50**, 1306 (1917).
(91) MEYER AND ROHMER: Ber. **33**, 250 (1900).
(92) MEYER AND STILLICH: Ber. **35**, 739 (1902).
(93) MILLER AND WAGNER: J. Am. Chem. Soc. **54**, 3698 (1932).
(93a) MILLS, HARRIS, AND LAMBOURNE: J. Chem. Soc. **119**, 1294 (1921).
(94) MÖHLAU: Ber. **31**, 2250 (1898).
(95) NASTYUKOV AND KRONEBERG: J. Russ. Phys.-Chem. Soc. **44**, 1200 (1912); Chem. Abstracts **7**, 983 (1913).
(96) NIEMENTOWSKI AND ORZECHOWSKI: Ber. **35**, 3898 (1902).
(97) ODDO AND TOGNACCHINI: Gazz. chim. ital. **52**, **II**, 347 (1922).
(97a) PANAJATOW: Ber. **20**, 32 (1887).
(98) PLÖCHL: Ber. **21**, 2117 (1888).
(99) PONZIO: Gazz. chim. ital. **24**, **II**, 280 (1894).
(100) PRATESI: Gazz. chim. ital. **14**, 351 (1884).
(101) PULVERMACHER: Ber. **25**, 2762 (1892).
(102) RAMEAU: Rec. trav. chim. **57**, 194 (1938).
(103) REINKING, DEHNEL, AND LABHARDT: Ber. **38**, 1077 (1905).
(104) ROHDE: Ber. **20**, 1911 (1887).
(105) RÜGHEIMER: Ber. **39**, 1653 (1906).
(106) SCHIFF: Ann. Chem. Pharm. Suppl. **3**, 343 (1864).
(107) SCHIFF: Ann. **140**, 93 (1867).
(108) SCHIFF: Ber. **25**, 1936 (1892).
(109) SCHIFF AND AMATO: Gazz. chim. ital. **1**, **I**, 376 (1871).
(110) SCHOLTZ: Ber. **32**, 2256 (1899).
(111) SCHOLTZ AND JAROSS: Ber. **34**, 1504 (1901).
(112) SCHWABBAUER: Ber. **35**, 413 (1902).
(113) SHIBATO AND OKUYAMA: Bull. Chem. Soc. Japan **11**, 122 (1936).
(114) SIMONS: J. Am. Chem. Soc. **59**, 518 (1937).
(115) SPIELMAN: J. Am. Chem. Soc. **57**, 583 (1935).
(116) STENHOUSE: Ann. **156**, 199 (1870).
(117) STRAIN: J. Am. Chem. Soc. **54**, 1221 (1932).

(118) Strache: Ber. **21**, 2361 (1888).
(119) Tollens: Ber. **17**, 653 (1884).
(120) Tollens and Mayer: Ber. **21**, 1570 (1888).
(121) Troeger: J. prakt. Chem. [2]**36**, 225 (1887).
(122) Tschitschibabin: J. Russ. Phys.-Chem. Soc. **37**, 1229 (1905).
(123) Tschitschibabin: J. prakt. Chem. [2]**107**, 122, 129, 132 (1924).
(124) Tschitschibabin and Aparina: J. prakt. Chem. [2]**107**, 138, 145, 154 (1924).
(125) Tschitschibabin and Moschkin: J. prakt. Chem. [2]**107**, 109 (1924).
(126) van Alphen: Rec. trav. chim. **54**, 93 (1935); **55**, 412, 669 (1936).
(127) van Alphen: Rec. trav. chim. **57**, 271 (1938).
(128) Veer: Rec. trav. chim. **57**, 994 (1938).
(129) von Braun: Ber. **41**, 2145 (1908).
(130) Von Miller: Ber. **25**, 2072 (1892).
(131) Von Miller and Kinkelin: Ber. **20**, 1916 (1887).
(132) Von Miller and Plöchl: Ber. **27**, 1296 (1894).
(132a) Von Miller and Plöchl: Ber. **29**, 1462 (1896).
(133) Von Miller and Plöchl with Eckstein: Ber. **25**, 2029 (1892).
(134) Von Miller and Plöchl with Eibner: Ber. **25**, 2041 (1892).
(135) Von Miller and Plöchl with Friedlander: Ber. **25**, 2049 (1892).
(136) Von Miller and Plöchl with Hofer: Ber. **25**, 2028 (1892).
(137) Von Miller and Plöchl with Krempelhuber: Ber. **25**, 2050 (1892).
(138) Von Miller and Plöchl with Lettermayer: Ber. **25**, 2039 (1892).
(139) Von Miller and Plöchl with Senders: Ber. **25**, 2033 (1892).
(140) Von Miller and Rohde: Ber. **25**, 2053 (1892).
(141) Wagner: J. Am. Chem. Soc. **55**, 724 (1933).
(142) Wagner: J. Am. Chem. Soc. **56**, 1944 (1934).
(143) Wagner: J. Am. Chem. Soc. **57**, 1296 (1935).
(144) Wagner: J. Org. Chem. **2**, 157 (1937).
(145) Wagner and Eisner: J. Am. Chem. Soc. **59**, 879 (1937).
(146) Wallach: Ber. **4**, 668 (1871).
(147) Wallach: Ann. **173**, 274 (1874).
(148) Walther and Bamberg: J. prakt. Chem. [2]**71**, 153 (1905).
(149) Walther and Bamberg: J. prakt. Chem. [2]**73**, 209 (1906).
(150) Walther and Kausch: J. prakt. Chem. [2]**56**, 111 (1897).
(151) Weil: Ber. **27**, 3317 (1894).
(152) Wellington and Tollens: Ber. **18**, 3298 (1885).
(153) Werner: J. Chem. Soc. **111**, 844 (1917).
(154) Wheeler: J. Am. Chem. Soc. **30**, 136 (1908).
(155) Wheeler and Jordan: J. Am. Chem. Soc. **31**, 937 (1909).
(156) Wheeler and Weller: J. Am. Chem. Soc. **24**, 1063 (1902).
(157) Winans and Adkins: J. Am. Chem. Soc. **55**, 2051 (1933).
(158) Wohl and Wohlberg: Ber. **34**, 1922 (1901).
(159) Wyckoff and Corey: Z. Krist. **89**, 462 (1934).
(160) Young and Wagner: J. Am. Chem. Soc. **59**, 854 (1937).
(161) Zaunschirm: Ann. **245**, 281 (1888).
(162) Zincke and Muhlhausen: Ber. **38**, 3824 (1905).

Recent patents

(163) I. G. Farbenindustrie, A. G.: British patent 474,034.
(164) I. G. Farbenindustrie, A. G.: British patent 474,601.
(165) I. G. Farbenindustrie, A. G.: French patent 817,153.
(166) Sutter (To Société pour l'industrie chimique á Bâle): U. S. patent 2,097,109.
(167) I. G. Farbenindustrie, A. G.: German patent 667,542.

THEORIES OF OPTICAL ROTATORY POWER

WALTER J. KAUZMANN, JOHN E. WALTER, AND HENRY EYRING

Frick Chemical Laboratory, Princeton University, Princeton, New Jersey

Received December 26, 1939

CONTENTS

I. INTRODUCTION

In an optically active molecule there are always two influences at work which render the molecule capable of rotating the plane of polarization of plane polarized light. The first of these is the purely structural one which permits the molecule to exist in non-superimposable mirror image forms (enantiomorphs); the second is that which makes possible the interaction of a given enantiomorphic form of the molecule with a plane polarized light wave so as to rotate its plane of polarization.

The structural influences which make possible the existence of enantiomorphic forms are simple and diverse. The first to be recognized was the asymmetric carbon atom, but there are many other factors besides the occurrence of such an atom in a molecule which may operate to give rise to non-superimposable mirror image forms. Thus, it is conceivable for

such a highly symmetrical molecule as ethane to occur in conformations which, if they could be isolated, would undoubtedly be found to be optically active (e.g., a conformation resulting from a rotation of one methyl group relative to the other from either the staggered or opposed form). The only feature which distinguishes those mechanisms for the production of enantiomorphs which the chemist has been able to uncover from those which are theoretically possible but which have not yet been proven operative is the fact that some enantiomorphs racemize much more readily than others. As the technique of the chemist becomes more and more refined, it will undoubtedly be found possible to resolve an ever-widening variety of types of substances into enantiomorphic forms.

In contrast to the simplicity of the structural conditions leading to optical activity in a substance, the mechanism of the interaction of such a substance with a light wave causing the plane of polarization to be rotated has resisted such complete interpretation. Thus, although it is easy to see why secondary butyl alcohol exists in two optically active forms, it is quite another matter to explain why it should have a molecular rotation of 10.3° in sodium D light at 20°C. in the absence of a solvent, and why this rotation should be different in different solvents and at different temperatures and wave lengths.

The structural conditions leading to optically active molecules have been adequately treated elsewhere, so need concern us no further here. It is the purpose of this paper to consider the mechanism of the interaction of an optically active molecule with a light wave from the standpoint of recent quantum-mechanical theories of rotatory power. It will be shown that these theories are capable of accounting for the observed orders of magnitude of optical rotations, and that, by showing how the rotatory power depends on molecular structure, they provide a powerful means of studying such structures under a wide variety of experimental conditions. They will also be found to be capable of supplying information concerning the nature of forces which act between molecules.

II. The Physical Basis of Optical Rotatory Power

A. Source of Optical Rotation; the Molecular Rotatory Parameter; Rotatory Strengths

The general theory of optical rotatory power has been presented in complete detail in the review article by Condon (8). We here summarize briefly the essential steps in the derivation of the fundamental equations, referring the mathematically inclined reader to Condon's paper for the justification of certain statements here given without proof.

The electromagnetic field in a region free from real charges or real currents is completely specified by Maxwell's equations:

$$\operatorname{div} \mathbf{D} = 0 \qquad \operatorname{div} \mathbf{B} = 0 \tag{1}$$

$$\operatorname{curl} \mathbf{E} = -\frac{1}{c}\frac{\partial}{\partial t}\mathbf{B} \qquad \operatorname{curl} \mathbf{H} = \frac{1}{c}\frac{\partial}{\partial t}\mathbf{D}$$

$$\mathbf{D} = \mathbf{E} + 4\pi\mathbf{P} \qquad \mathbf{B} = \mathbf{H} + 4\pi\mathbf{M}$$

(Heavy type indicates vector quantities.) For an isotropic medium which is not optically active, we may write

$$\mathbf{D} = \epsilon\mathbf{E} \qquad \mathbf{B} = \mu\mathbf{H} \tag{2}$$

where ϵ is the dielectric constant and μ is the magnetic permeability. The solution of equations 1 then represents electromagnetic waves propagated with a velocity

$$v = \frac{c}{\sqrt{\epsilon\mu}}$$

where c is the velocity of light. For all cases in which we shall be interested, μ is very nearly unity, therefore $v = c/\sqrt{\epsilon}$. If n is the index of refraction of the medium, we have the familiar result $n^2 = \left(\frac{c}{v}\right)^2 = \epsilon$. For media which are optically active, equations 2 are no longer complete. If N_1 is the number of molecules per cubic centimeter, we have $\mathbf{P} = N_1\mathbf{p}$, $M = N_1\mathbf{m}$, where $\mathbf{p}$ and $\mathbf{m}$ are the induced electric and magnetic moments per molecule. A theory which enables us to calculate $\mathbf{p}$ and $\mathbf{m}$ in terms of the structure of the individual molecule will thus give us the correct expressions of the form 2, and hence a complete theory of the propagation of electromagnetic waves through the medium in question. The quantum-mechanical calculations of Rosenfeld (34) showed that the electric and magnetic moments induced in a molecule by the perturbing electromagnetic field are given by the equations:

$$\mathbf{p}_a = \alpha_a\mathbf{E}' - \frac{\beta_a}{c}\frac{\partial}{\partial t}\mathbf{H}' + \gamma_a\mathbf{H}' \tag{3}$$

$$\mathbf{m}_a = \qquad \frac{\beta_a}{c}\frac{\partial}{\partial t}\mathbf{E}' + \gamma_a\mathbf{E}'$$

where $\mathbf{E}'$ is the effective electric field on the molecule,
$\mathbf{H}'$ is the effective magnetic field on the molecule,

$$\alpha_a = \frac{2}{3h} \sum_b \frac{\nu_{ab} \,|\, (a/p/b) \,|^2}{\nu_{ab}^2 - \nu^2}$$

$$\beta_a = \frac{c}{3\pi h} \sum_b \frac{\mathrm{Im}\ \{(a/\mathbf{p}/b) \cdot (b/\mathbf{m}/a)\}}{\nu_{ab}^2 - \nu^2}$$

$$\gamma_a = \frac{2}{3h} \sum_b \frac{\nu_{ab}\ \mathrm{Re}\ \{(a/\mathbf{p}/b) \cdot (b/\mathbf{m}/a)\}}{\nu_{ab}^2 - \nu^2}$$

The expressions for α_a, β_a, γ_a have been averaged over all orientations of the molecule with respect to the field, assuming all orientations to be equally probable. $\mathbf{p}_a$, $\mathbf{m}_a$, α_a, β_a, γ_a refer to the molecule in the state a; $(a/\mathbf{p}/b)$ and $(b/\mathbf{m}/a)$ are the matrix components of the electric and magnetic moments connecting states a and b; ν_{ab} is the frequency of the transition between states a and b; Im and Re mean that we are to take the imaginary or real part of the scalar product $(a/\mathbf{p}/a) \cdot (b/\mathbf{m}/a)$, which will be in general complex.

The terms in γ_a will produce only a very small second-order effect on the optical rotation (8), and are neglected in the following discussion. For an isotropic medium, we use the Lorentz field

$$\mathbf{E}' = \mathbf{E} + \frac{4\pi N_1}{3} \mathbf{p}$$

This relation seems to hold quite well even for liquids in which the molecular distribution would not necessarily be random. Since the intensity of magnetization M is quite small for non-magnetic media, we set $\mathbf{H}' = \mathbf{H}$. We shall need to consider only those states a which are available to the molecule at ordinary temperatures. The molecule may exist in various rotational-vibrational levels, or various configurations due to free rotation of the groups of the molecule. The average induced moments may then be written

$$\mathbf{p} = \alpha\left(\mathbf{E} + \frac{4\pi N_1}{3} \mathbf{p}\right) - \frac{\beta}{c} \frac{\partial}{\partial t} \mathbf{H} \tag{4}$$

$$\mathbf{m} = \frac{\beta}{c} \frac{\partial}{\partial t}\left(\mathbf{E} + \frac{4\pi N_1}{3} \mathbf{p}\right)$$

where $\alpha = \sum_a \rho_a \alpha_a$, $\beta = \sum_a \rho_a \beta_a$, and ρ_a is the probability that the molecule be in the state a of a certain conformation. α is the polarizability and β the molecular rotatory parameter of the molecule.

Use of equations 4 gives for **D** and **B** the relations

$$\begin{aligned}\mathbf{D} &= \mathbf{E} + 4\pi\mathbf{P} = \mathbf{E} + 4\pi N_1\mathbf{p}\\ &= \left(\frac{3 + 8\pi N_1\alpha}{3 - 4\pi N_1\alpha}\right)\mathbf{E} - \frac{12\pi N_1\beta/c}{3 - 4\pi N_1\alpha}\frac{\partial}{\partial t}\mathbf{H} \qquad (5)\\ \mathbf{B} &= \mathbf{H} + 4\pi\mathbf{M} = \mathbf{H} + 4\pi N_1\mathbf{m}\\ &= \mathbf{H} + \frac{12\pi N_1\beta/c}{(3 - 4\pi N_1\alpha)}\frac{\partial}{\partial t}\mathbf{E} \qquad \text{(to the first order in } \beta)\end{aligned}$$

By identifying the coefficient of **E** in equations 2 and 5 and indicating the coefficient of $\frac{\partial}{\partial t}$ **H** by g we have

$$\frac{\epsilon - 1}{\epsilon + 2} = \frac{4\pi N_1\alpha}{3} \qquad (6)$$

$$g = 4\pi N_1\beta/c\,\frac{\epsilon + 2}{3} \qquad (7)$$

Introducing ϵ and g, the equations 5 take the form:

$$\mathbf{D} = \epsilon\mathbf{E} - g\frac{\partial}{\partial t}\mathbf{H} \qquad \mathbf{B} = \mathbf{H} + g\frac{\partial}{\partial t}\mathbf{E} \qquad (8)$$

Equation 6 is the familiar relation between dielectric constant and the molecular polarizability. The optical rotatory power of the medium will be found to be a function of g.

For a plane wave moving along the z-axis, the general solution for **D** may be written in the form

$$\mathbf{D} = \mathbf{D}_0 e^{i\psi} \qquad (9)$$

where

$$\psi = 2\pi\nu\left(t - \frac{z}{v}\right) = 2\pi\nu\left(t - \frac{nz}{c}\right)$$

with similar expressions for **B**, **E**, and **H**. We shall be interested only in the real part of **D** in equation 9, and in the following equations **D** refers only to the real part of the general solution (equation 9).

D can be proved from equation 1 to be perpendicular to the direction of propagation, and so may be written in terms of its components as $\mathbf{D} = \mathbf{i}D_x + \mathbf{j}D_y$.

D for right circularly polarized light (i.e., **D** rotating clockwise as viewed by an observer looking along the $-z$-axis) must be of the form

$$\mathbf{D}_r = D(\mathbf{i}\cos\psi_r - \mathbf{j}\sin\psi_r)$$

Similarly for left circularly polarized light

$$\mathbf{D}_l = D(\mathbf{i} \cos \psi_l + \mathbf{j} \sin \psi_l)$$

Now we see that equation 9 will reduce to the above equations if we take the correct expression for the vector amplitude, i.e., if we take

$$\mathbf{D}_r = \mathrm{Re}\{D(\mathbf{i} + i\mathbf{j})e^{i\psi_r}\}$$
$$\mathbf{D}_l = \mathrm{Re}\{D(\mathbf{i} - i\mathbf{j})e^{i\psi_l}\} \tag{10}$$

where D is the scalar amplitude. The solution of Maxwell's equations subject to the conditions 8 shows that the indices of refraction are different for the two types of circular polarization (8). The results are

$$n_r = \epsilon^{1/2} - 2\pi\nu g \tag{11}$$
$$n_l = \epsilon^{1/2} + 2\pi\nu g$$

We therefore write

$$\psi_r = \psi_0 + \delta \qquad \psi_l = \psi_0 - \delta \tag{12}$$

where

$$\psi_0\left(= 2\pi\nu\left(t - \frac{nz}{c}\right)\right)$$

is the phase of a wave propagated with the mean index of refraction $n = \epsilon^{1/2}$ and

$$\delta = 4\pi^2\nu^2 g \frac{z}{c}$$

The superposition of the right and left circularly polarized waves leads to a plane polarized wave. If we substitute equations 12 in 10 and add, we have

$$\mathbf{D} = \mathbf{D}_r + \mathbf{D}_l = 2\mathrm{D} \cos \psi_0 \,(\mathbf{i} \cos \delta - \mathbf{j} \sin \delta)$$

For $\delta = 0$, we have a plane polarized wave, with the electric induction vector in the x direction. For $\delta > 0$, the plane of polarization has been rotated through an angle δ in the clockwise sense as viewed in the z direction. The rotation of the plane of polarization in radians per centimeter is therefore

$$\varphi' = \frac{\delta}{z} = \frac{4\pi^2\nu^2}{c} g = \left(\frac{2\pi}{\lambda}\right)^2 cg \tag{13}$$

By comparison with equation 11, we may also write

$$\varphi' = \frac{\pi}{\lambda}(n_l - n_r) \tag{14}$$

i.e., a medium is dextrorotatory if the refractive index for left circularly polarized light is greater than that for right circularly polarized light. Experimentally, the rotation φ is usually expressed in degrees per decimeter, or

$$\varphi = \frac{1800}{\pi} \varphi'$$

The specific rotation, $[\alpha]$, is defined as $[\alpha] = \varphi/\rho$, where ρ is the density of the active material. The molecular rotation, $[M]$, is usually defined as

$$[M] = [\alpha] \frac{M}{100}$$

where M is the molecular weight of the active material. Using the above relations, and expressing g in terms of the molecular rotatory parameter β by equation 7, we have

$$[M] = \frac{288\pi^2 N}{\lambda^2} \frac{n^2 + 2}{3} \beta \tag{15}$$

$$[M]_{\mathrm{D}} = 4.93 \times 10^{35} \frac{n^2 + 2}{3} \beta_{\mathrm{D}} \text{ for sodium D light}$$

Substituting for β

$$[M] = \frac{96\pi N}{hc} \frac{n^2 + 2}{3} \sum_{a_i} \rho_{a_i} \sum_{b_i} \frac{R_{b_i a_i} \nu^2}{\nu_{b_i a_i}^2 - \nu^2} \tag{15a}$$

where

$$R_{b_i a_i} = \mathrm{Im}\ \{(a_i/\mathrm{p}/b_i) \cdot (b_i/\mathrm{m}/a_i)\}$$

$R_{b_i a_i}$ is called the rotatory strength of the transition $a_i \rightarrow b_i$. ρ_{a_i} is the probability that the molecule lies in a particular electronic state with definite relative positions for the atoms. Any particular relative position of the atoms in a molecule we call a *conformation*. The electronic state a_i includes all those corresponding translational, vibrational, and rotational states which have no appreciable influence on the optical rotation, so that a_i is actually made up of a set of states. If only one electronic level is accessible to the molecules in thermal equilibrium, as is usually the case, we may speak of the set of states a_i as making up a conformation i, so that ρ_{a_i} is now equal to the probability of the molecule's occurring in a conformation i.

In order to obtain some idea of the orders of magnitude of the quantities involved in equation 15, we may substitute values for a typical active substance, *sec*-butyl alcohol. Here $[M]_{\mathrm{D}} = 10.3°$ and $n_{\mathrm{D}} = 1.397$, so that we find $\beta = 1.59 \times 10^{-35}$.

Equation 15a corresponds to the dispersion formula in general use in expressing the variation of rotation with wave length:

$$[M] = \sum_i \frac{A_i}{\lambda^2 - \lambda_i^2} \tag{15b}$$

In going from equation 15a to equation 15b, however, the effect of variation of refractive index with wave length has been neglected. The factor $(n^2 + 2)/3$ remains nearly constant throughout the visible for most substances, but in careful work and in work extending into the ultraviolet its variation should be taken into account.

B. SYMMETRY PROPERTIES OF ROTATORY STRENGTHS

The operators $\mathbf{p}$ and $\mathbf{m}$ expressed in cartesian coördinates are

$$\mathbf{p} = e(\mathbf{i}x + \mathbf{j}y + \mathbf{k}z)$$

$$\mathbf{m} = \frac{eh}{4\pi mci}\left\{\mathbf{i}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right) + \mathbf{j}\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) + \mathbf{k}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\right\} \tag{16}$$

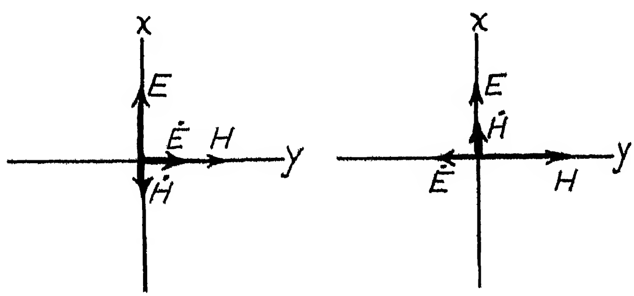
If the molecule has a center of symmetry, we may classify the states of the molecule as odd or even, according as the wave function for a given state changes sign or retains the same sign when we reflect the function through the center of symmetry, i.e., replace each coördinate by its negative. Since the operator $\mathbf{p}$ changes sign upon reflection, we have a non-vanishing value of $(a/p/b)$ only between odd and even states. The operator $\mathbf{m}$ does not change sign upon reflection, hence we have a non-vanishing value of $(b/\mathbf{m}/a)$ only between two odd or two even states. The scalar product $(a/\mathbf{p}/b)\cdot(b/\mathbf{m}/a)$ will therefore be identically zero for all states a and b, and the optical rotation will vanish. If the molecule has a plane of symmetry, we may again classify the wave functions as even or odd with respect to reflection in this plane. Thus when there is a plane of symmetry there is no pair of states a and b for which the x components of $(a/\mathbf{p}/b)$ and of $(b/\mathbf{m}/a)$ are both different from zero. The same is true of the y and z components. A fundamental requirement for optical activity is thus that the molecule possess neither a plane nor a center of symmetry.

We now consider the values of $(a/\mathbf{p}/b)\cdot(b/\mathbf{m}/a)$ for a given molecule and its mirror image. A molecule may be transformed into its mirror image by reflection of its coördinates in any plane, which is equivalent to changing from a right-handed to a left-handed coördinate system. If we reflect a molecule in the x–y plane, the new value of $(a/\mathbf{p}/b)\cdot(b/\mathbf{m}/a)$ for the molecule is obtained if we replace z by $-z$ in the wave functions a and b in the corresponding expression for the old molecule. If we change z to $-z$ in both eigenfunctions and operators, the product of the matrix

components will remain unchanged, as the values of the integrals involved, being pure numbers, are independent of the particular coördinate system in which they are evaluated. As we see from equation 16, $\mathbf{p} \cdot \mathbf{m}$ changes sign upon replacing z by $-z$; therefore $(a/\mathbf{p}/b) \cdot (b/\mathbf{m}/a)$ must change sign if we replace z by $-z$ in the eigenfunctions only. An unsymmetrical molecule and its mirror image must therefore have equal and opposite optical rotations. If a molecule is the same as its mirror image, its optical rotation must vanish.

C. A QUALITATIVE DISCUSSION OF THE ORIGIN OF OPTICAL ROTATION

The preceding discussion can be made more explicit by giving a simple physical picture of the meaning of certain of the equations. Equation 4 states that there is a contribution to the induced electric moment $\mathbf{p}$ proportional to the time rate of change of the magnetic field, and a contribution to the induced magnetic moment $\mathbf{m}$ proportional to the time rate of change of the electric field. We may visualize how such contributions could arise if we assume that the potential field in an unsymmetrical molecule is such that the electrons are not free to move in the direction of applied forces, but are constrained to move along a spiral path with a component of motion normal to the impressed force. If we have a circularly polarized wave moving along the z-axis (z-axis pointing below the plane of the paper), the electric and magnetic vectors and their time derivatives will at some instant have the configuration shown in figure 1 (the xz-plane being the plane of polarization):



I Right circular polarization

II Left circular polarization

FIG. 1. Electric and magnetic vectors and their time derivatives

Let us focus our attention on configuration I, and assume that the electron is forced to move along a right-handed spiral about the $+$ x-axis. The electric dipole moment will arise largely from the displacement of the electron in the $-x$ direction under the influence of the field E. But by

Lenz's law, the changing magnetic field will tend to move the electron in the right-handed direction about the $+x$-axis around a circle in the plane perpendicular to $\dot{H}$. Since the electron is forced to move along the spiral path, it will be displaced in the $+x$ direction, and thus we will have a contribution to the electric moment in the direction of E, and proportional to $\dot{H}$. This proportionality constant is related to β in equation 4; and, as we see from our simple example, it depends upon the pitch of the spiral and hence upon the degree to which the potential field is unsymmetric. The changing electric field will tend to move the electron along the $-y$-axis. The spiral forces the electron to move in a circular path, thus producing a magnetic moment proportional to $\dot{E}$. Considering the induced electric moment only, we have for configuration I, $p_{x_{\mathrm{I}}} = aE - b\dot{H}$. For configuration II, since E is in the same direction as before but $\dot{H}$ is in the opposite direction, we have $p_{x_{\mathrm{II}}} = aE + b\dot{H}$, where the constants a and b will be the same in the two cases. The amplitude of the vibration of the electron will thus be greater for left than for right circularly polarized light; the former therefore loses more energy to the molecule than does the latter. This means that $n_l > n_r$ and a medium composed of such "molecules" would be dextrorotatory.

We shall now consider equation 15 in terms of the above model. The only quantities in this equation whose magnitudes depend upon the nature of the model are ν_{ba} and $\mathbf{p}_{ba}\cdot\mathbf{m}_{ba}$. ν_{ba} is the frequency of a certain type of motion executed by the model. Classically, $\mathbf{p}_{ab}\cdot\mathbf{m}_{ba}$ is $\mathbf{P}\cdot\mathbf{M}$, where $\mathbf{P}$ and $\mathbf{M}$ are the vector amplitudes of the variable electric and magnetic moments associated with this type of motion. For an electron moving periodically back and forth along a spiral, $\mathbf{M}$ is directed essentially along the axis of the spiral, while $\mathbf{P}$ has a component along this direction, so that $\mathbf{P}\cdot\mathbf{M}$ will not vanish and there is a non-vanishing β. The greater the pitch of the spiral, the greater will be the component of $\mathbf{P}$ in the direction of $\mathbf{M}$, hence the greater will be the optical rotatory power of the model. If, on the other hand, the pitch of the spiral is zero, so that the electron moves in a circle, then $\mathbf{P}$ and $\mathbf{M}$ will be perpendicular, and the model will possess no optical activity. Thus we see that there is a direct relationship between the magnitude of $\mathbf{P}\cdot\mathbf{M}$ and b in the relation $p_x = aE \pm b\dot{H}$ given above.

D. ROTATORY STRENGTHS, LINE STRENGTHS, AND ANISOTROPY FACTORS

It can be shown (8) that for each state a

$$\sum_b R_{ba} = 0$$

Therefore all the quantities R_{ba} cannot have the same sign. We see from equation 15 that for $\nu = \infty$

$$[M] = \text{const.} \sum_b R_{ba} = 0$$

also for $\nu = 0$, $[M] = 0$. Optical rotatory power is thus a property that tends to zero for very large and for very small frequencies.

The polarizability can be written as

$$\alpha = \frac{2}{3h} \sum_b \frac{\nu_{ba} S_{ba}}{\nu_{ba}^2 - \nu^2} \tag{17}$$

where $S_{ba} = |(a/\mathbf{p}/b)|^2$ = line strength of the transition $a \to b$ or

$$\alpha = \frac{e^2}{4\pi^2 m} \sum_b \frac{f_{ba}}{\nu_{ba}^2 - \nu^2} \tag{18}$$

$$f_{ba} = \frac{8\pi^2 m}{3e^2 h} \nu_{ba} S_{ba} = \text{oscillator strength of the transition}$$

For the oscillator strengths, we have the relation that for each state a,

$$\sum_b f_{ba} = n$$

where n is the number of electrons in the molecule (17, 40). Since S_{ba} is always positive, while R_{ba} may be positive or negative, we see that optical rotatory power is a small effect in comparison to ordinary refraction.

Kuhn (19) defines the anisotropy factor g_{ab} for the transition $a \to b$. For gases ($n = 1$) the definition leads to the expression

$$\frac{(n_l - n_r)_{a \to b}}{(n - 1)_{a \to b}} = \frac{\nu}{\nu_{ba}} g_{ba} \tag{19}$$

In terms of the quantum-mechanical quantities defined above,

$$g_{ba} = 4 \frac{R_{ba}}{S_{ba}}$$

which is valid for condensed phases as well as for gases. The anisotropy factor is thus the number by which the contribution of a given transition to the polarizability must be multiplied to obtain the contribution to the optical rotatory power. In most cases, the wave functions of a molecule will be approximately either even or odd, hence for a given transition either $(a/\mathbf{p}/b)$ or $(b/\mathbf{m}/a)$ may be large, but in general both will not be large. Therefore strong absorption bands will have small anisotropy factors, and weak bands may have large anisotropy factors, with the result that the rotational strengths of both strong and weak absorption bands may be of the same order of magnitude.

E. CIRCULAR DICHROISM

An optically active medium also has different absorption coefficients for right and left circularly polarized light. The quantum-mechanical treat-

ment has been presented in detail by Condon, Altar, and Eyring (9). For a given transition, it is found that

$$\frac{\epsilon_l - \epsilon_r}{\epsilon} = 4\frac{\nu}{\nu_{ba}}\frac{R_{ba}}{S_{ba}} = \frac{\nu}{\nu_{ba}} g_{ba} \qquad (20)$$

where ϵ is the mean absorption transition probability, and ϵ_r and ϵ_l are the absorption transition probabilities for right and left circularly polarized light. This unequal absorption of left and right circularly polarized light is known as circular dichroism, and is experimentally measured by determining the ellipticity of an originally plane polarized beam, the ellipticity per unit length being

$$\theta = \frac{\pi}{\lambda}(\epsilon_l - \epsilon_r)$$

Measurements of the absorption and ellipticity for a given band will give us the anisotropy factor g_{ba} for this band, and thus enable us to determine rotatory strengths from line strengths.

The ellipticity is a maximum where the absorption is a maximum, i.e., in the absorption band. Taking

$$\epsilon = \frac{8\pi^3}{3h^2c} S_{ba}$$

and using equation 15 for R_{ba} in equation 20 we obtain:

$$[M]_{ba} = \frac{3hN}{\pi^2}\frac{\nu\nu_{ba}(\epsilon_l - \epsilon_r)(n^2 + 2)}{\nu_{ba}^2 - \nu^2}$$

This gives Natanson's rule that if a medium absorbs left circularly polarized light more strongly than right, then the partial rotation of the given transition is positive for $\nu < \nu_{ab}$.

The expressions given above for α and β are of course not correct for $\nu = \nu_{ab}$. In this case, as Condon points out, analogy with the quantum-mechanical theory of dispersion gives for the molecular rotation:

$$[M] = \frac{96\pi N}{hc}\frac{n^2 + 2}{3}\sum_a \rho_a \sum_b \frac{\nu^2(\nu_{ba}^2 - \nu^2)R_{ba}}{[(\nu_{ba}^2 - \nu^2)^2 + \nu^2\Gamma_{0ba}^2]}$$

The ellipticity per unit length then has the form:

$$\theta = \frac{\pi}{\lambda}(\epsilon_l - \epsilon_r) = \frac{16\pi^2 N_1}{3hc}\sum_a \rho_a \sum_b \frac{\nu^3\Gamma_{0ba}R_{ba}}{[(\nu_{ba}^2 - \nu^2)^2 + \nu^2\Gamma_{0ba}]}$$

Here Γ_0 (the half-width of the line) measures the strength of the damping factor. The change of sign of the partial contribution, plus the ellipticity produced in the band, is known as the Cotton effect; the general behavior of these quantities in the neighborhood of a band is shown diagrammatically in figure 2.

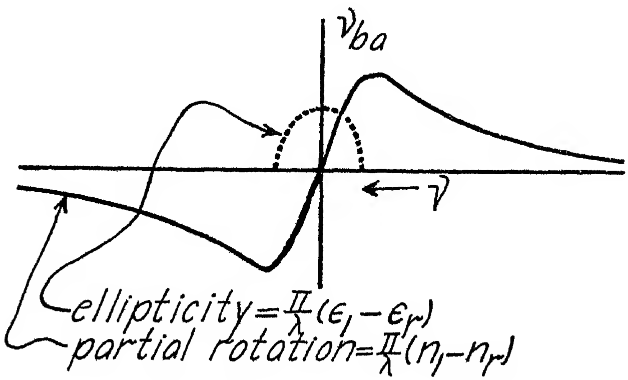


FIG. 2. General behavior of the ellipticity and the partial rotation in the neighborhood of a band

F. PARTIAL ROTATIONS, CHROMOPHORIC GROUPS, AND VICINAL ACTION

We have seen that the molecular rotatory power may be expressed as

$$[M] = \text{const} \sum_b \frac{R_{ba}\nu^2}{\nu_{ba}^2 - \nu^2}$$

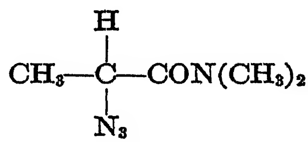
where each term in the sum represents the contribution of a single transition. Only electronic transitions will be of importance in producing optical rotation, as we see from the manner in which the mass enters into the expressions for magnetic moment in equations 16. Thus transitions involving changes in nuclear vibrations or rotations will be less effective than electronic transitions in contributing to optical activity by a factor of the order of the ratio of nuclear to electronic masses. For transparent substances, the electronic transitions with which we will be concerned will lie in the ultraviolet. Because of the factor ν_{ba}^2 in the denominator, transitions in the far ultraviolet will be of less importance than transitions of comparable rotatory strengths in the near ultraviolet, and in addition the contributions from the far ultraviolet absorption lines will tend to cancel out, since the several values for R_{ba} will not all be of the same sign. In many cases, therefore, the resultant optical rotation will be due largely to the one or two absorption bands nearest the visible. If the dispersion

data are accurate enough, and especially if the dispersion has been measured through the band, the constants for the bands nearest the visible may be fairly accurately determined, any residual rotation being expressed by a term representing the resultant of the remaining bands. It must be emphasized that in the cases where the empirical dispersion data are expressed by a single term, the constants R_{ba} and ν_{ba} determined by this data will in general represent some sort of average of the constants of several bands.

If the partial rotation of a single transition can be separated from the residual rotation, it is frequently found that this transition is localized in a particular group of the molecule, called the chromophoric group. The remaining groups of the molecule affect this partial rotation only through their vicinal effect on the chromophoric group, that is, they produce the dissymmetry necessary to make the transition in the chromophoric group optically active.

The general discussion given above may be made more specific by several examples given by Kuhn (19, 20, 22).

In



there is an absorption band at 2900 Å. which can be associated with a transition in the N_3 group. The oscillator strength of this band can be determined by integrating the absorption over the band.

$$\int \epsilon_\nu \, d\nu = N_1 f \frac{\pi e^2}{mc} \left(\frac{n^2 + 2}{3} \right)^2$$

The value found for this band is $f = 5 \times 10^{-4}$. Since $\Sigma f \sim 50$, we see that the N_3 band at 2900 Å. accounts for about 10^{-5} of the total absorption and polarizability. The anisotropy factor for this band is $g = 2 \times 10^{-2}$, and the band contributes about 20 per cent of the observed rotation in the visible.

In camphor, the band at 3000 Å. can be associated with a transition in the C═O group, as this band appears in all aldehydes and ketones. Measurements of circular dichroism and absorption show that this band is actually two superposed electronic bands, the one nearest the visible being optically active, the other inactive. The active band has an f value of 2×10^{-4}, the inactive band an f value several times greater. This active band represents only about the 3×10^{-5} part of the total absorption, but its partial rotation accounts for the greater part of the observed optical rotation.

III. Theories of Optical Rotatory Power

A. THE COUPLED-OSCILLATOR THEORY

With the derivation of equation 15 by means of quantum mechanics, the problem of optical activity can be said to be solved in principle. Any modern theory seeking to relate observed optical rotations to molecular structure should have this equation as its starting point. Two such theories have been proposed in recent years,—the polarizability theory of Kirkwood and the one-electron theory of Condon, Altar, and Eyring. Considerable success has been obtained, however, in the interpretation of

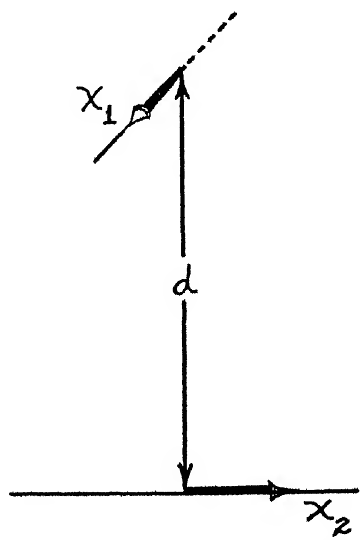


FIG. 3. Simplest system of coupled oscillators which gives optical activity

experimental data, at least in a qualitative sense, by the use of the coupled-oscillator theory of Born (5), Oseen (31), and Kuhn (18), especially in the form developed and applied by Kuhn. Although this theory was originally formulated in classical terms, it has since been derived from equation 15 for Kuhn's simple model by Condon (8).

For this simplest system of coupled oscillators (see figure 3) which gives optical activity, we consider that there are in the molecule two oscillators of charge e_1, e_2, and mass m_1, m_2, which in the absence of interactions will vibrate at right angles to one another with frequencies ν_1', ν_2', the positions of equilibrium being separated by a distance d.

If now there is a coupling between the oscillators, that is, if the potential energy has the form $V = 1/2\, k_1x_1^2 + 1/2\, k_1x_2^2 + k_{12}x_1x_2$, the resulting mo-

tion can be resolved into two normal vibrations by a transformation to a new set of axes which are related to the old set by a rotation through an angle α, which will be small if the coupling constant k_{12} is small. The normal frequencies, ν_1, ν_2, will be nearly equal to the original frequencies ν_1', ν_2' in the case of small coupling. In the classical treatment it is shown that the energy transferred to these normal vibrations is different for right and left circularly polarized light, resulting in differing indices of refraction for the two types of circular polarization, and thus in optical rotation. In the quantum-mechanical treatment, the rotatory strength, R_{ba}, is calculated by the use of harmonic oscillator wave functions. Either treatment of this model leads to the result, for N_1 molecules per unit volume, oriented at random:

$$\alpha = \frac{2\pi N_1}{3}\frac{n^2+2}{3}\frac{d}{\lambda^2}\sin\alpha\cos\alpha\,\frac{e_1 e_2}{(m_1 m_2)^{1/2}}\left\{\frac{1}{\nu_1^2-\nu^2}-\frac{1}{\nu_2^2-\nu^2}\right\} \quad (21)$$

The presence of two non-parallel oscillators in a molecule will thus give rise to an optical rotation. Since the coupling forces between these oscillators may be expected to be small, both oscillators will maintain their characteristic frequencies, but by their mutual interaction they will become optically active. Oscillator No. 1 is said by its *vicinal action* to cause an *induced anisotropy* in oscillator No. 2 and *vice versa*. It is in this way that Kuhn is able to account for the optical rotation associated with an absorption band of a symmetrical group such as the carbonyl in an unsymmetrical molecule such as camphor.

Inspection of equation 21 shows immediately that the sum rule is obeyed, since the rotatory strength associated with ν_1 is the negative of that associated with ν_2.

Since the details of this important theory and especially its application in interpreting experimental data have been thoroughly covered in a review article by Kuhn (21), we shall not devote so much space to it in the following pages as to the more recent theories.

B. POLARIZABILITY THEORY OF KIRKWOOD

Equation 15a has been applied to the calculation of optical rotations by Kirkwood (15), who was able to transform it into an expression for rotatory power in terms of the polarizabilities and anisotropies of the groups in the molecule. The underlying principle is that the electrons of a molecule may be considered as assignable to definite groups between which there is no exchange. (This is nearly true of all electrons in lower states except the small fraction involved in the bonding of one group to another.) Furthermore, each transition which a molecule can undergo is to a first approximation localized in a distinct group. (This is justified by the fact

that, for instance, the presence of a carbonyl group in a molecule is always accompanied by an absorption band in the neighborhood of 2900 Å., and similarly for other groups.)

The first of these assumptions is introduced by breaking up the electric moment and magnetic moment terms in equation 15a into sums over groups:

$$\mathbf{p}_{ab} = \sum_k \mathbf{p}_{ab}^{(k)}$$

$$\mathbf{m}_{ba} = \sum_k \left(\mathbf{R}_k \times \mathbf{P}_{ba}^{(k)} + \frac{2mc}{e} \mathbf{m}_{ba}^{(k)} \right)$$

where the superscript k refers to the k^{th} group, $\mathbf{R}_k$ is the radius vector of the center of gravity of the k^{th} group relative to some fixed point in the molecule, $\mathbf{P}^{(k)}$ is the total electronic momentum operator of group k, $\mathbf{p}^{(k)}$ is the electric moment operator of group k, and $\mathbf{m}^{(k)}$ is the magnetic moment operator of the k^{th} group relative to its center of gravity. Substituting in equation 15a,

$$\beta = \frac{c}{3\pi h} \sum_{a,b} \left\{ e^{-\frac{\epsilon_a}{kT}} \frac{\operatorname{Im}\,(\mathbf{p}_{ab} \cdot \mathbf{m}_{ba})}{\nu_{ba}^2 - \nu^2} \right\} = \beta^{(0)} + \beta^{(1)} + \sum_k \beta_k \qquad (22)$$

where

$$\beta^{(0)} = \frac{1}{3h} \sum_{i>k} \sum_a e^{\frac{-\epsilon_a}{kT}} \sum_{b>a} \frac{\nu_{ba}}{\nu_{ba}^2 - \nu^2} \operatorname{Re}(\mathbf{R}_{ik} \cdot (\mathbf{p}_{ab}^{(i)} \times \mathbf{p}_{ba}^{(k)})) \qquad (22a)$$

$$\beta^{(1)} = \frac{c}{3\pi h} \sum_{i>k} \sum_a e^{\frac{-\epsilon_a}{kT}} \sum_{b>a} \frac{\operatorname{Im}\,(\mathbf{p}_{ab}^{(k)} \cdot \mathbf{m}_{ba}^{(i)})}{\nu_{ba}^2 - \nu^2} \qquad (22b)$$

$$\beta_k = \frac{c}{3\pi h} \sum_k \sum_a e^{\frac{-\epsilon_a}{kT}} \sum_{b>a} \frac{\operatorname{Im}\,(\mathbf{p}_{ab}^{(k)} \cdot \mathbf{m}_{ba}^{(k)})}{\nu_{ba}^2 - \nu^2} \qquad (22c)$$

$\mathbf{R}_{ik} = \mathbf{R}_k - \mathbf{R}_i$ = vector from center of gravity of i to that of k. β_k is assumed negligible to the approximation considered here for symmetrical groups, it being the contribution of the k^{th} group alone to the optical rotation parameter. $\beta^{(1)}$, the contribution due to coupling of magnetic moment on one group with an electric moment on another, is neglected as probably small, although it deserves further consideration. $\beta^{(0)}$ therefore remains as the major source of the rotatory power; it is similar in origin to the coupling of oscillators in the theories of Born and Kuhn.

The second approximation (that of localization of transitions within definite groups) is now introduced in order to calculate $\beta^{(0)}$. A set of zero-order eigenfunctions is set up, characterized by quantum numbers assignable to each of the groups of the molecule. First-order perturbation theory is applied, using as the perturbing potential, V, the expression for

thei nteraction of two dipoles, $\mathbf{p}_i$ and $\mathbf{p}_j$, at a distance $\mathbf{R}_{ij}$ apart, where this distance is large compared with the separation of charge in the dipoles:

$$V = \frac{p_i p_j}{R_{ij}^3} (\cos\theta - 3\cos\chi\cos\psi) \tag{23}$$

where χ is the angle between $\mathbf{p}_i$ and $\mathbf{R}_{ij}$, ψ is the angle between $\mathbf{p}_j$ and $\mathbf{R}_{ij}$, and θ is the angle between $\mathbf{p}_i$ and $\mathbf{p}_j$.

The perturbed eigenfunctions are now used to calculate the $\mathbf{p}_{ba}^{(k)}$, and it is found that by a series of transformations it is possible to replace the product

$$\mathbf{R}_{ik} \cdot (\mathbf{p}_{ab}^{(i)} \times \mathbf{p}_{ba}^{(k)})$$

by an expression involving the polarizabilities, anisotropies, and relative orientations of the principal axes of the groups i and k.

The de Mallemann theory of optical rotatory power results from an application of second-order perturbation theory, while the Boys theory would come out of a third-order calculation. Whereas the first-order calculation results in an expression involving simultaneous interactions of pairs of groups, the second- and third-order calculations result in expressions involving simultaneous interactions of three and four groups, respectively.

The final expression obtained by Kirkwood is:

$$\beta^{(0)} = -\frac{1}{6} \sum_{i>k} \sum_{r,s} \alpha_{rr}^{(i)} \alpha_{ss}^{(k)} (\mathbf{b}_r^{(i)} \cdot \mathbf{T}_{ik} \cdot \mathbf{b}_s^{(k)}) \mathbf{R}_{ik} \cdot (\mathbf{b}_r^{(i)} \times \mathbf{b}_s^{(k)}) \tag{24}$$

where

$$\mathbf{T}_{ik} = \frac{1}{R_{ik}^3} \left(1 - 3\,\frac{\mathbf{R}_{ik}\mathbf{R}_{ik}}{R_{ik}^2}\right)$$

$\alpha_{rr}^{(i)}$ are the three polarizabilities along the principal axes $\mathbf{b}_r^{(i)}$ of group i, and similarly for $\alpha_{ss}^{(k)}$ and $\mathbf{b}_s^{(k)}$.

If the groups in the molecule are cylindrically symmetrical (two of the $\alpha_{rr}^{(i)}$ equal), this may be simplified to:

$$\beta^{(0)} = -\frac{1}{6} \sum_{i>k} \alpha_i \alpha_k \delta_i \delta_k (\mathbf{b}_i \cdot \mathbf{T}_{ik} \cdot \mathbf{b}_k)(\mathbf{R}_{ik} \cdot \mathbf{b}_i \times \mathbf{b}_k) \tag{25}$$

where α_i and α_k are the mean polarizabilities of the interacting groups i and k, and δ_i and δ_k are their anisotropies:

$$\delta_i = \frac{\alpha_{11}^{(i)} - \alpha_{22}^{(i)}}{\alpha_i}$$

where $\alpha_{11}^{(i)}$ is the polarizability along axis of cylindrical symmetry, $\mathbf{b}_i$, and similarly for δ_k. The values of the polarizabilities α_i involved here

can be found from the molecular refractions of the groups present. The anisotropies δ_i may be estimated from the depolarization of light by simple derivatives containing the group in question, and also from the Kerr constant of such substances. The other quantities, $\mathbf{b}_i$, $\mathbf{R}_{ik}$, etc., are determined by a knowledge of the structure of the molecule in question. The vectors $\mathbf{b}_i$ are, of course, of unit length. Unfortunately there may sometimes be some doubt as to the direction of some of the $\mathbf{b}$'s due to our inadequate knowledge of the origin of optical anisotropy.

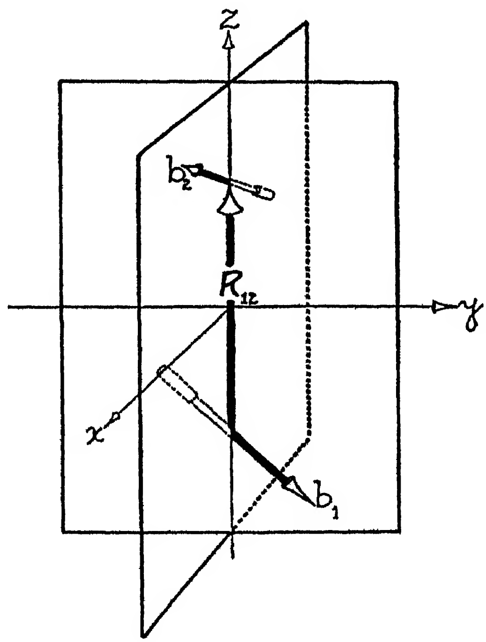


FIG. 4

The above treatment bears a very close resemblance to the calculation of van der Waals forces, and indeed it may be said that the interactions which are taken into account are those which give rise to those forces.

There are some indications that this theory is not able to deal adequately with all factors observed in connection with optical rotatory power. This is particularly true of the rotations associated with weak absorption bands and this is a typical defect of all polarizability theories, as was pointed out by Lowry and Allsopp (26). Thus, in camphor, for example, it is known that a major portion of the observed rotation is associated with the transition at 2950 Å. in the carbonyl group (20). Yet this band is known to contribute very little to the polarizability of a carbonyl-contain-

ing molecule (26). This fact can only be brought into conformity with the theory if the anistropy ratio, δ, were to change markedly with wave length without altering the polarizability, which is unlikely. It seems probable that in this theory the optical activity associated with weak bands is contained in the terms $\beta^{(1)}$ and $\beta^{(k)}$, since such bands probably obtain the large values of their anisotropy factors from a large value for the magnetic moment of their transitions (see later).

A further limitation to the theory is found in the inadequacy of the potential, V, of equation 23. This expression is only true for dipoles which are separated by distances large compared with the separation of charge within the dipoles. In the case under consideration, however, this condition is far from being met, especially when perturbations involving excited states are involved, since the eigenfunctions for these extend out over a considerable region of space and even overlap considerably most of the nearby groups. Of the similar calculation of the van der Waals forces between helium atoms, it is known (29) that the term in the potential given by equation 23 becomes seriously inadequate at a distance of 3 Å. This distance would be increased when interactions between excited states are concerned. Therefore, it would seem that the theory cannot be expected to give good results for more compact molecules where interactions between groups closer than, say, 4 or 5 Å. give rise to the optical activity.

It is instructive to calculate the optical rotation of a very simple model (figure 4) on this theory. Let there be two cylindrically symmetric groups whose principal axes have the following directions:

$$\mathbf{b}_1 = \frac{1}{\sqrt{2}}(\mathbf{j} - \mathbf{k})$$

$$\mathbf{b}_2 = \frac{1}{\sqrt{2}}(\mathbf{i} + \mathbf{k})$$

The vector from the center of group No. 1 to the center of group No. 2 is

$$\mathbf{R}_{12} = d\mathbf{k}$$

so that d is the distance between them. Using these quantities in equation 25 we find that

$$\mathbf{b}_1 \cdot \mathbf{T}_{12} \cdot \mathbf{b}_2 = \frac{1}{d^3}$$

$$R_{12} \cdot (b_1 \times b_2) = \frac{1}{2}(\mathbf{i} - \mathbf{j} - \mathbf{k}) \cdot d\mathbf{k} = -\frac{d}{2}$$

$$\beta^{(0)} = +\frac{\alpha_1 \alpha_2 \delta_1 \delta_2}{12d^2}$$

If the principal axes of the hydroxyl groups of hydrogen peroxide have the orientations postulated here (and they are probably not far from it (33)), we may calculate the optical rotation of one of the optically active forms of this substance. Taking $\alpha = 10.3 \times 10^{-25}$ cc., $\delta = 0.35$, $d = 1.25$ Å., and refractive index $= 1.414$, we find that

$$\beta^{(0)} = +6.8 \times 10^{-35}$$

and

$$[M]_{\mathrm{D}} = +45°$$

on this model.

If $\delta_1 = \delta_2 = 1$ we have the groups acting as linear oscillators, and the model takes on the aspect of that of Kuhn, except that more explicit account is taken here of the physical nature of the coupling of the oscillators.

Since on this theory the optical rotation depends on the products of two polarizabilities, and since the polarizability depends upon the frequency, as follows,

$$\alpha = \sum_i \frac{k_i}{\nu_i^2 - \nu^2}$$

the effect on the optical rotation of changing the frequency of the light should be given by

$$[M] = \sum_{i>j} \frac{k_{ij}\nu^2}{(\nu_i^2 - \nu^2)(\nu_j^2 - \nu^2)} = \sum_{i>j} \nu^2 k'_{ij} \left\{\frac{1}{\nu_i^2 - \nu^2} - \frac{1}{\nu_j^2 - \nu^2}\right\} \quad (26)$$

A test of this relationship to find if it agrees with the observed behavior of rotatory dispersion is at present beyond the ability of experimental technique.

C. THE ONE-ELECTRON THEORY OF OPTICAL ACTIVITY

In applying equation 15 to actual molecules, the problem is to select a model which has the twin properties of tractability and accuracy in essential features. In a mathematical sense the problem is completely solved, but this is not a particularly interesting remark to one interested in obtaining numbers to compare with experiment. An inspection of experimental dispersion curves for optical rotatory power reveals, as equation 15 already suggests, a striking correlation between absorption and optical rotation. The absorption by a molecule in the visible and ultraviolet is due to electronic transitions, so that electronic transitions must be the central feature of any model that is to treat the dispersion satisfactorily.

Atomic spectra have been successfully interpreted as arising principally

from one-electron transitions (10). The same procedure is usually adopted in interpreting most molecular transitions.[1] We thus expect that rotatory power can be treated by the same procedure. Our next problem is, therefore, to find the intial and final states of these one-electron transitions. We note that the transitions which a molecule can undergo are generally each characteristic of one of the groups in the molecule. Thus hydroxyl-containing compounds have an absorption band at $\sim$ 1800 Å., while compounds with an unconjugated C═C bond absorb at $\sim$ 2150 Å. In speaking of the optical activity associated with these transitions, we refer to the group primarily involved in the transition as the *chromophoric group* and the forces which act on it to make its transition optically active as *vicinal actions*. Therefore, in calculating the optical rotation we must first find out the natures of the electronic states of these groups. Insofar as previous descriptions of transitions define these states for the important chromophoric groups, the procedure here is straightforward. The perturbation of these electronic states by neighboring radicals leads to altered states which give transitions which are optically active provided the system as a whole has no plane or center of symmetry.

The general procedure is as follows: The eigenfunctions for two non-degenerate states of a chromophore which are acted on by a perturbation V can, using the usual perturbation theory, be expressed as:

$$\begin{aligned} \psi_a &= \psi_a^0 + \sum_{i \neq a} c_{ia}\psi_i^0 \\ \psi_b &= \psi_b^0 + \sum_{j \neq b} c_{jb}\psi_j^0 \end{aligned} \tag{27}$$

where

$$c_{ia} = \frac{\int \psi_i^0 V \psi_a^0 \, \mathrm{d}\tau}{E_a^0 - E_i^0} \qquad c_{jb} = \frac{\int \psi_j^0 V \psi_b^0 \, \mathrm{d}\tau}{E_b^0 - E_j^0}$$

Using these new functions in calculating R_{ba} in equation 15, we find

$$\begin{aligned} R_{ba} = \mathbf{p}_{ab}\cdot\mathbf{m}_{ba} &+ \sum_i c_{ia}(\mathbf{p}_{ib}\cdot\mathbf{m}_{ba} + \mathbf{p}_{ab}\cdot\mathbf{m}_{bi}) + \sum_j c_{jb}(\mathbf{p}_{aj}\cdot\mathbf{m}_{ba} + \mathbf{p}_{ab}\cdot\mathbf{m}_{ja}) \\ &+ \sum \text{(terms involving products of two } c\text{'s)} \\ &+ \sum \text{(terms involving products of three } c\text{'s)} + \cdots \end{aligned} \tag{28}$$

where

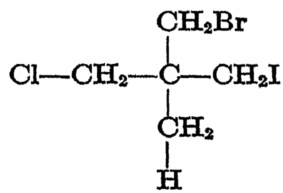
$$\mathbf{p}_{ab} = \int \psi_a^0 \mathbf{p} \psi_b^0 \, \mathrm{d}\tau; \quad \mathbf{p}_{ai} = \int \psi_a^0 \mathbf{p} \psi_i^0 \, \mathrm{d}\tau; \quad \mathbf{m}_{ab} = \int \psi_a^0 \mathbf{m} \psi_b^0 \, \mathrm{d}\tau; \text{ etc.}$$

For the unperturbed states ψ_a^0 and ψ_b^0 which are usually taken there is no optically active transition, so that $\mathbf{p}_{ab}\cdot\mathbf{m}_{ba}$ is zero. Furthermore, a

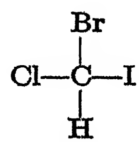
[1] See, for example, papers by R. S. Mulliken.

good many of the products $\mathbf{p} \cdot \mathbf{m}$ are either zero or very small, so that in practice only a few of the coefficients c_{ia} and c_{ib} need be calculated.

We shall see that the coefficients, c, are made up of sums of contributions from the interactions of individual groups with each chromophore. Therefore, the optical rotation can be regarded as made up of the sums of interactions between groups taken two at a time, plus the sums over products of two of these interactions (i.e., sums over interactions involving three groups at a time), and so on. The first sum gives what we shall call the first-order contributions to the optical activity, the second sum gives the second-order contributions, etc. Since the interactions are always small (all c's < 1) the first-order effects will be the largest in general, while the second-order effects will be next largest, etc. And, since, as we shall see, vicinal actions fall off with increasing distance, the interactions involving the groups which are farther off will show a greater tendency to give rise primarily to first-order contributions. Thus the second-order contribution in



will be much less than that in



If we are to use equation 28 in calculating the optical activity of a given substance, we must decide (*a*) what eigenfunctions are to be used in describing the unperturbed states of the electrons in the chromophoric group and (*b*) what forces from the surrounding radicals perturb these states.

(*a*) Whether we choose to describe a state of the chromophoric electron in terms of a linear combination of eigenfunctions for the three-dimensional oscillator, or in terms of the hydrogen-like functions, or in terms of any other complete orthogonal set is entirely a matter of convenience. This is true whether the electron spends its time moving about one center or moving about many centers. Only the two types of function mentioned, however, have been applied to the problem of calculating optical rotations on the one-electron model, and of these the hydrogen-like functions are the closer to reality, that is, they require fewer terms to give a good picture of the true state of affairs.

(*b*) The perturbing field V can usually be treated as being made up of a number of central force fields, V_j, arising from the various groups and atoms surrounding the given chromophore, the subscript j referring to the j^{th} such atom or group. The contribution to any c_{ia} by a V_j involves the calculation of the two-center integral

$$\int \psi_i^0 V_j \psi_a^0 \, d\tau$$

For hydrogen-like eigenfunctions which are not spherically symmetrical, this integral will depend upon the direction of the center j from the center of the eigenfunction with respect to a set of axes defined by the eigenfunction. Now it is generally true that a transformation can be applied

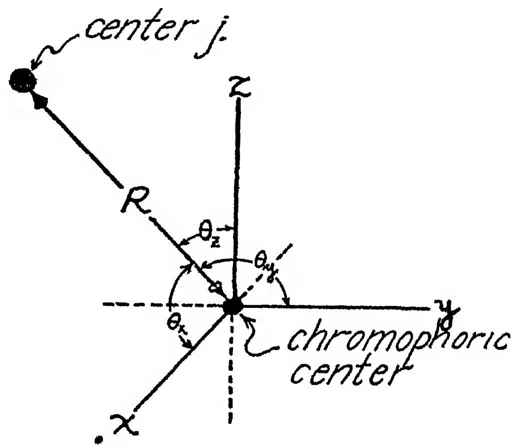


FIG. 5

which will permit the integral to be expressed as a sum of the products of two types of functions (13); one of the types depends upon the distance, R, between the perturbing center and the center of the chromophoric group under consideration, while the other type depends upon the angles which the line **R** joining the two groups makes with the chosen coördinate axes of the eigenfunction of the chromophoric group. This may be represented as follows (the angles θ are as given in figure 5):

$$\int \psi_i^0 V_j \psi_a^0 \, d\tau = F(R) f(\theta_x \theta_y \theta_z) + G(R) g(\theta_x \theta_y \theta_z) + \cdots \tag{29}$$

where the sum usually contains a small number of terms, and where the forms of the functions F, f, G, g, $\cdots$ depend upon the nature of ψ_a^0, ψ_i^0, and V_j.

In the event that the potential V_j is not spherically symmetric about the center j but depends upon the angles φ_x, φ_y, φ_z, which the line R

makes with a set of coördinate axes in the center j, the integral may be expressed as

$$\int \psi_i^0 V_j \psi_a^0 \, d\tau = F(R) f(\theta_x \theta_y \theta_z) \Phi(\varphi_x \varphi_y \varphi_z) + G(R) g(\theta_x \theta_y \theta_z) \Gamma(\varphi_x \varphi_y \varphi_z) \quad (30)$$

The functions f, g, $\cdots$ are the angular dependence functions of the chromophoric group; Φ, Γ, $\cdots$ are the angular dependence functions of the vicinal group. These functions are usually relatively simple expressions involving products of sines and cosines of angles. The functions F, G, $\cdots$ are the radial dependence functions of the vicinal action between the vicinal and chromophoric groups and are usually complicated functions of R best expressed by a graphical plot.

When the perturbing fields V_j are sufficiently small to make the second-order effects mentioned before negligible, the optical rotation is given in terms of these functions by:

$$[M] = \frac{96\pi N}{hc} \frac{n^2 + 2}{3} \sum_a \rho_a \sum_b R_{ba} \nu^2 / \nu_{ba}^2 \quad \nu^2 \quad (31)$$

$$R_{ba} = \sum_{j,k} \frac{(Ff\Phi + Gg\Gamma + \cdots)_{jk}}{E_a^0 - E_k^0} (\mathbf{p}_{kb} \cdot \mathbf{m}_{ba} + \mathbf{p}_{ab} \cdot \mathbf{m}_{bk})$$

$$+ \sum_{j,l} \frac{(F'f'\Phi' + G'g'\Gamma' + \cdots)_{jl}}{E_b^0 - E_l^0} (\mathbf{P}_{al} \cdot \mathbf{m}_{ba} + \mathbf{P}_{ab} \cdot \mathbf{m}_{la})$$

The first sum in R_{ba} refers to the effect of perturbations on the lower state, while the second sum refers to the effect on the upper state of the given transition. The primes and subscripts on F, g, Φ, etc., refer to the fact that these must be calculated for each perturbation j and for each perturbing level k and l acting on each of the levels a and b of the transition.

It is seen from equation 31 that insofar as the first-order contribution to the optical rotation is concerned, the rotation is additive in the vicinal actions of the perturbing groups in the sense that the introduction of a new vicinal group adds a quantity to the partial rotation of a given chromophore which is independent of the number, nature, and arrangement of the other vicinal groups.

The problem of calculating optical rotations by the one-electron theory therefore resolves itself into a calculation of the radial and angular dependence functions for the various types of fields and states which we find in a given problem.

An example of these functions and their use in calculating the optical rotation will be given later. Meanwhile, we must investigate the forces which can be expected to act as perturbations in giving rise to optical rotatory power.

Perturbations which can act as vicinal forces

(*1*) *Dipole forces* may be treated in two ways: they may be considered as arising from a separation of charges, the effect of each charge being calculated as a separate spherically symmetrical perturbation, or we may use the expression for the potential of an electron in the field of a dipole of moment μ, the axis of which makes an angle θ with the line of length R to the point at which the potential is calculated:

$$V = \frac{\mu e}{R^2} \cos \theta \tag{32}$$

Here the angular dependence function of the perturbing field is cos θ if the distance between dipole and chromophore is large.

That the fields from dipoles in a dissymmetric molecule may give rise to appreciable optical rotations was shown by Condon, Altar, and Eyring (9), who were able to account for the order of magnitude of the partial rotation of the nitrite band in phenylmethylcarbinol nitrite by using harmonic oscillator eigenfunctions and the accepted values for the bond moments occurring in the molecule. A further indication that dipole fields may affect the optical rotation in an important way is found in the work of Betti, Rule, and Beckmann and Cohen (see later). The latter workers, however, interpret the dipole effect as being due to a distortion of the molecular framework.

That dipoles are not sufficient to account for the observed rotations in all cases, however, is shown by the work of Condon, Altar, and Eyring on *sec*-butyl alcohol. Following the same procedure here as for phenylmethylcarbinol nitrite a rotation too small by two powers of ten was obtained.

(*2*) *The fields of ions and ionic charges* in the neighborhood of a dissymmetric molecule should produce a marked effect on the rotation, since they are more intense than dipole fields. These have the form

$$V = \frac{Ze^2}{DR} \tag{33}$$

where D is the dielectric constant and Ze the ionic charge. Here, of course, the angular dependence function of the perturbing group is independent of angle, hence is equal to unity.

Linear fields should also be capable of producing optical activity. Here

$$V = -zeF \tag{34}$$

where z is the distance measured in the direction of the field F.

(*3*) Since the electronic clouds about a nucleus are spread out over a finite region, the potential in the neighborhood of even a neutral atom is

not zero. Therefore, if the electron cloud of the chromophoric group overlaps those of the surrounding atoms, a perturbation from this source will result. This is especially likely to occur when the chromophoric electron is in an excited state. An example is the potential at a distance r from a neutral hydrogen atom in its lowest state:

$$V_{\mathrm{H}} = -\frac{e^2}{a_0} e^{-\frac{2r}{a_0}}\left(1 + \frac{a_0}{r}\right) \tag{35}$$

where a_0, the Bohr radius of a hydrogen atom, equals 0.529 Å. These potentials are usually spherically symmetric or nearly so about the perturbing atom.

Calculations by Gorin, Walter, and Eyring have shown that these forces are adequate to account for the observed rotations of *sec*-butyl alcohol and some of the sugars (12, 13).

(*4*) When there is any overlapping of the electronic orbits, it follows that the Pauli principle must be considered; this gives rise to the so-called "*exclusion forces*" or "*exchange repulsions*" similar in origin to the exchange forces in a valence bond. In order to evaluate the relative importance of these forces and the classical coulombic forces which were considered directly above and with which they will always be associated, we can compare their values in a molecule such as hydrogen for which they are well known. In making the comparison, however, we must remember that the coulombic repulsion between the hydrogen nuclei is included in the energy of a hydrogen molecule, and that this interaction is of no importance in determining the optical rotation arising from the interaction of two groups, since here only those interactions which involve electrons are important. This internuclear repulsion amounts to 440 kg-cal. for a nuclear separation of 0.76 Å. The energy of H_2 is 103 kg-cal., of which about 10 per cent, or 10 kg-cal., is the net coulombic energy (including nuclear repulsions) and 90 per cent, or 90 kg-cal., is exchange energy. Now, if the energy of repulsion of nuclei is 440 kg-cal. and the net coulombic energy is 10 kg-cal., the attractive energy of the electrons interacting with the nuclei less the repulsive energy of the electrons interacting with each other must be 450 kg-cal., or five times the exchange energy. (Of course, in this calculation our point of zero energy was taken as the two separated hydrogen atoms.) Furthermore, in bonds between atoms other than hydrogens the exchange energy usually makes up a smaller proportion of the bond energy than in the case of H_2, so that the importance of exchange forces relative to coulombic forces in causing optical activity may be even less.

It is also observed that in the excited states of hydrogen, triplet states (i.e., those in which exchange forces oppose molecule formation) are very

nearly as stable as the corresponding singlet states (i.e., those in which the exchange forces favor molecule formation). For example:

INTER-NUCLEAR DISTANCE	STATE	CONFIGURATION	ENERGY ABOVE NORMAL H_2	DIFFERENCE	ENERGY OF DISSOCIATION
1.08 Å.	$^1\Sigma_g^+$ $^3\Sigma_g$	1s 3d	111810 cm.$^{-1}$ = 319.665 kg-cal. 111860 cm.$^{-1}$ = 319.808	0.143 kg-cal.	About 60 kg-cal. for all
~1.06 Å.	$^1\Sigma_g^+$ $^3\Sigma_g^+$	1s 4d	117404 cm.$^{-1}$ = 335.658 117522 cm.$^{-1}$ = 335.996	0.338 kg-cal.	
~1.07 Å.	$^1\Pi_g$ $^3\Pi_g$	1s 4d	117574 cm.$^{-1}$ = 336.144 117656 cm.$^{-1}$ = 336.379	0.235 kg-cal.	

The data are from reference 39.

This would indicate that where excited states are involved, exclusion forces are still less important than they are between atoms in normal states. This is the more significant because the overlapping of orbits is greatest where excited states are involved.

It is noted that in general the ratio of exchange to coulombic energies is approximately equal to the square of the overlap integral, and this is usually a small quantity.

(5) *van der Waals forces.* It is only valid to speak of van der Waals forces between groups when the groups are rather far apart so that their electronic clouds do not overlap, since otherwise the assumptions underlying the quantum-mechanical calculation leading to their formulation break down. Therefore, in small compact molecules like *sec*-butyl alcohol, and more especially in considering the perturbations on excited states, care must be exercised in applying this concept.

In the one-electron theory, explicit account is not taken of the van der Waals forces between groups. This neglect is probably not serious, especially for more compact molecules, since in the calculation of the effects of the forces due to incomplete screening of atoms, a procedure is automatically adopted similar to that which Hartree has used so successfully in studying atomic energies; here the chromophoric electron is assumed to be moving in a static field due to the other electrons and protons of the system. From the success of the Hartree method in giving wave functions which in turn give good values for energy levels and charge distributions in atoms, we can assume that the same method will give the correct wave functions for optical activity, especially as the molecules being considered become more and more compact and also when the wave functions of excited states are calculated. From the failure of the method to account for

van der Waals forces at larger separations, it evidently tends to become less true at greater distances between chromophoric and vicinal groups. But because of the breakdown of the assumptions underlying the study of van der Waals forces at intermediate distances, it is probably the best available approximation to the true state of affairs in these regions.

It is to be noted again that the forces considered by Kirkwood when he formulated his theory (see elsewhere) are essentially van der Waals forces, so that the formula for optical rotation which he obtains should tend to be more valid as groups are more widely separated from one another.

The large optical rotations associated with weak absorption bands are understandable in the one-electron theory, since weak bands (i.e., transitions with small electric moment) may have large magnetic moments. Thus, the carbonyl band at 2950 Å. is believed (30) to be due to a transition of a non-bonding $2p_y$ electron on the oxygen to either a (Z_C-Z_O) or an (X_C-X_O) antibonding orbital. The situation may be roughly approximated by description of the transition in terms of hydrogen-like orbitals located on the oxygen atom. Then the transition is between a $2p_y$ state and a $2p_x$ or $2p_z$ state, either of which would give rise to a large magnetic moment, which could then couple with the small electric moment associated with the band to give rise to optical rotation.

In general, in the one-electron theory, although the sum rule, $\sum_b R_{ba} = 0$, still holds true, it does not hold true because of pairwise cancellations such as occur in the Kuhn and Kirkwood theories. Therefore, it will not be possible to write the dispersion in the form of equation 26.

Owing to the inadequacy of the eigenfunctions used in practice in calculating the optical rotation by the one-electron theory, one is restricted to giving at most signs of the partial rotations associated with given transitions and their relative orders of magnitude. The theory, however, allows us to form a new and broader concept of vicinal action which we shall find to be very useful in interpreting some of the experimental results concerning optical rotation.

Calculation of optical rotation on the one-electron model

In order to show that the one-electron theory can account for the observed order of magnitude of optical rotations in general, and also to illustrate the type of procedure which is followed in calculating a rotation on this theory, we now consider an explicit example. A simple transition for our purpose, as expressed in atomic, hydrogen-like orbitals, is one between a $2p_y$ state and a $2p_z$ state. Such a transition bears some resemblance to the transition which gives rise to the 2950 Å. absorption band of the

carbonyl group, which according to Mulliken (30) is between a $2p_y$ non-bonding electron on the carbonyl oxygen and a (Z_C-Z_O) antibonding orbital between the carbon and the oxygen[2] (the coördinate axes referred to here are as given in figure 6). Therefore, we shall assume that a transition, $2p_y \rightarrow 2p_z$, is occurring on the oxygen atom of a carbonyl group, and that this transition will give approximately, so far as the optical rotation is concerned, the behavior of an electron actually located in the carbonyl group.[3]

The simplest optically active carbonyl-containing compound is 3-methylcyclopentanone; because of its simplicity we shall calculate the partial rotation of its 2950 Å. absorption band on the above model. Since there are no appreciable dipoles or charges in the molecule which can perturb the carbonyl group, the major portion of this rotation probably comes from the incomplete screening of the nuclei of atoms surrounding the

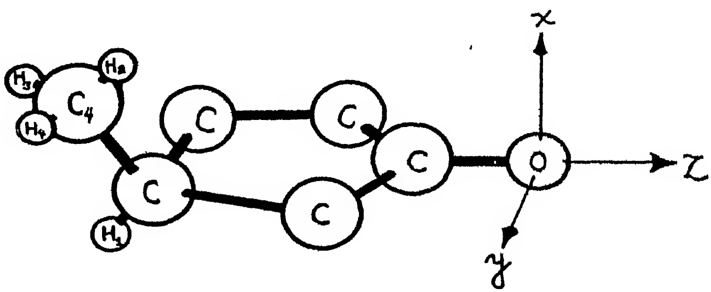


FIG. 6. Model of 3-methylcyclopentanone

carbonyl by their electrons, so that this type of vicinal action alone will be considered here.

When we apply equation 28, with $\psi_a^0 = 2p_y$ and $\psi_b^0 = 2p_z$, and employ only the approximation which gives the lowest-order contribution to the optical rotation, we must recognize that since the electron cloud of the excited state is more spread out in space than that of the lower state, the upper state will be perturbed to a greater degree by the incomplete screening of neighboring nuclei, and hence will give a larger contribution to the optical rotation arising from this source. (At the distances involved in the present example, the upper state overlaps the perturbing atoms to an extent about fifteen times greater than does the lower state for the eigenfunctions actually used.) Secondly, it is to be noted that the transi-

[2] It is possible that X_C-X_O is the upper state. This corresponds to a transition $2p_y \rightarrow 2p_x$ here, and an actual calculation shows that it would give a rotation which is equal but opposite in direction to that found for $2p_y \rightarrow 2p_z$.

[3] A more adequate treatment of the carbonyl will be made at a later date.

tion $2p_y \rightarrow 2p_z$ gives rise to a large magnetic moment in the x direction and no electric moment in any direction, while the only perturbing function with a principal quantum number of 3 or less with which $2p_y$ will give an electric moment component in this direction is $3d_{x+y}$. Although some of the possible perturbing functions with principal quantum numbers larger than 3 will also give an electric moment component in this direction, this component will be very small, since the matrix element $(2p_y/x/nd_{x+y})$ decreases rapidly with increasing n.

Thus we find that as far as the first-order and major contribution of the transition to the rotation is concerned, the transition is

$$a = \psi_{2p_y} \longrightarrow b = (\psi_{2p_z'} + C\psi_{3d_{x+y}})$$

The rotatory strength for the transition is, from equation 31,

$$\begin{aligned} R_{ba} &= \mathrm{Im}\ \{[\psi_{2p_y}/\mathbf{p}/(\psi_{2p_z'} + C\psi_{3d_{x+y}})]\cdot[(\psi_{2p_z'} + C\psi_{3d_{x+y}})/\mathbf{m}/\psi_{2p_y}]\} \\ &= C\ \mathrm{Im}\ \{(\psi_{2p_y}/ex/\psi_{3d_{x+y}})(\psi_{2p_z'}/m_x/\psi_{2p_y})\} \\ &= -\frac{e^2 h}{4\pi mc}(\psi_{2p_y}/x/\psi_{3d_{x+y}})(\psi_{2p_z'}/1/\psi_{2p_z})C \end{aligned}$$

since

$$m_x = \frac{eh}{4\pi mci}\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)$$

Now

$$\beta' = \frac{c}{3\pi h}\frac{R_{ba}}{\nu_{ba}^2 - \nu^2}$$

where β' is the contribution of the absorption band in question to the rotation of the plane of light of frequency ν. Taking the D line of sodium as the light employed ($\lambda = 5890$ Å., $\nu = 0.507 \times 10^{15}$ sec.$^{-1}$) and taking the absorption band at 2950 Å. ($\nu = 1.02 \times 10^{15}$ sec.$^{-1}$), it is found that

$$\beta' = -2.89 \times 10^{-24}(\psi_{2p_y}/x/\psi_{3d_{x+y}}))\psi_{2p_z'}/1/\psi_{2p_z})C$$

Using the usual hydrogen-like eigenfunctions:

$$\psi_{2p_y} = \frac{1}{4\sqrt{2\pi}}\left(\frac{Z_1}{a_0}\right)^{5/2}\left(e^{-\frac{Z_1 r}{2a_0}}\right)y$$

$$\psi_{2p_z'} = \frac{1}{4\sqrt{2\pi}}\left(\frac{Z_2}{a_0}\right)^{5/2}\left(e^{-\frac{Z_2 r}{2a_0}}\right)z$$

$$\psi_{3d_{x+y}} = \frac{\sqrt{2}}{81\sqrt{\pi}}\left(\frac{Z_2}{a_0}\right)\left(e^{-\frac{Z_2 r}{3a_0}}\right)xy$$

and setting $Z_1 = 1.8$ and $Z_2 = 1.5$ from the observed values of the ionization potential of a $2p_y$ electron in the carbonyl and the energy of the transition under consideration, it is found that

$$(\psi_{2p_y}/x/\psi_{3d_{x+y}}) = 1.077 \times 10^{-8} \text{ cm.}$$

$$(\psi_{2p_z'}/1/\psi_{2p_z}) = 0.98$$

Substituting these values in the expression for β', we find that

$$\beta' = -3.04 \times 10^{-32}\, C$$

It remains, therefore, to find C. Now,

$$C = \frac{\int \psi_{2p_z'} V \psi_{3d_{x+y}}\, d\tau}{E_{2p_z'} - E_{3d_{x+y}}}$$

and

$$E_{2p_z'} - E_{3d_{x+y}} = -\frac{e^2}{2a_0} 1.5^2 \left(\frac{1}{2^2} - \frac{1}{3^2}\right) = -\frac{5}{32}\frac{e^2}{a_0}$$

V is here the sum of the potentials due to the surrounding atoms. We are concerned only with potentials due to incomplete screening of the nuclei for hydrogen atoms and carbon atoms, for which, at a distance r from the respective nuclei,

$$V_H = -\frac{e^2}{a_0} e^{-\frac{2r}{a_0}} \left(1 + \frac{a_0}{r}\right) \tag{35}$$

$$V_C = -\frac{4e^2}{a_0} e^{-\frac{2.95r}{a_0}} \left(1.07 \left(\frac{r}{a_0}\right)^2 + 2.175 \left(\frac{r}{a_0}\right) + 2.213 + \frac{a_0}{r}\right) \tag{36}$$

The integrals $\int \psi_{2p_z'} V \psi_{3d_{x+y}} d\tau$ are best evaluated using elliptical coördinates, with the foci at the centers of the perturbing atom and the oxygen atom. A rotation of axes must be performed, however, which will bring the Z-axis into line with the line joining the centers of the two atoms. Because of the cylindrical symmetry about the new Z-axis of the potentials V_H and V_C and of the radial parts, ψ_{2p} and ψ_{3d}, of the eigenfunctions, the integrals of odd powers of x and y vanish. And again, because of this symmetry, it is possible to arrange the x- and y-axes so that an integral over any power of y is the same as that over the same power of x. Then the rotation of axes will be found to transform the product xyz (which arises from the angular factors z and xy of $\psi_{2p_z'}$ and $\psi_{3d_{x+y}}$) into $\gamma_x\gamma_y\gamma_z(Z'^3 - X'^2Z')$, where γ_x, γ_y, and γ_z are the direction cosines of the line joining the two atomic centers in the odd coördinate system, and where z' and x' are the new coördinates. Since the overlapping of the functions in the integral $\int \psi_{2p_z'} V \psi_{3d_{x+y}} d\tau$ will be greatest where z' is rather large and

x' rather small, we may neglect the term x'^2z' as compared with z'^3. Thus we have that

$$C = \gamma_x\gamma_y\gamma_z D$$

where

$$D = -\frac{32}{5}\frac{a_0}{e^2}\int \psi_{2p'} V \psi_{3d} z'^3 \, d\tau$$

and where $\psi_{2p'}$ and ψ_{3d} refer to the eigenfunctions without the angular factors z and xy.

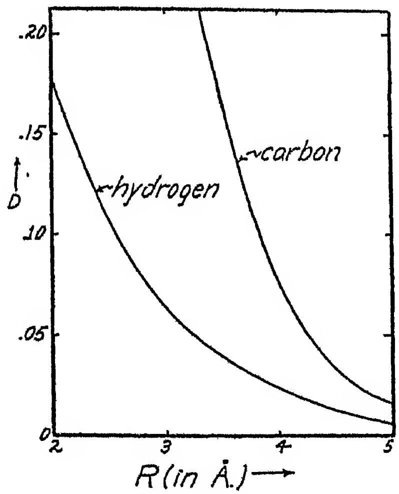


FIG. 7. Vicinal action of carbon and hydrogen atoms

For β' and $[M]'_D$, the contribution to the molecular rotation, we find

$$\beta' = -3.04 \times 10^{-32} \gamma_x\gamma_y\gamma_z D$$

$$[M]'_D = -15.1 \times 10^8 \frac{n^2 + 2}{3} \gamma_x\gamma_y\gamma_z D \tag{37}$$

Naturally, the value of D will depend upon the distance between the perturbing group and the chromophore. This dependence for V_H and V_C is given in figure 7. Referring to the terminology in a previous section (page 363), $\gamma_x\gamma_y\gamma_z$ is the angular dependence function of the carbonyl, while D, aside from the factor $-\frac{32}{5}\frac{a_0}{e^2}$, is the radial dependence function.

We now possess all the information necessary to calculate the partial rotation of the present approximation to the carbonyl group in 3-methyl-

cyclopentanone from a knowledge of the structure of the molecule. The refractive index here is 1.434. The other necessary quantities appearing in equation 37 are given in table 1. R in this table is the distance from the oxygen to the atom in question. All distances are in Ångström units. The calculated value of the partial molecular rotation measured in D light for the pseudocarbonyl transition in methylcyclopentanone is +59.2° for the configuration given in figure 6. The observed total molecular rotation of methylcyclopentanone is 130°. (The two values are not directly comparable.) Thus it is seen that the observed order of

TABLE 1

Coördinates of atoms in 3-methylcyclopentanone

	O	C_4	H_1	H_2	H_3	H_4
x	0	1.258	−0.889	2.148	1.258	1.258
y	0	1.293	1.140	0.923	0.943	2.325
z	0	−4.339	−4.130	−3.830	−5.071	−3.989
R		4.699	4.377	4.487	5.300	4.785
$\gamma_x = \frac{x}{R}$		+0.268	−0.203	+0.479	+0.237	+0.263
$\gamma_y = \frac{y}{R}$		+0.275	+0.261	+0.206	+0.178	+0.486
$\gamma_z = \frac{z}{R}$		−0.923	−0.944	−0.854	−0.957	−0.834
$\gamma_x\gamma_y\gamma_z$		−0.0662	+0.0480	−0.0841	−0.0404	−0.1065
D		+0.025	+0.016	+0.013	+0.002	+0.008
Contribution to $[M]'_D$		+33.8°	−15.7°	+22.2°	+1.6°	+17.3°

magnitude of the rotation can be accounted for sufficiently well on this model with only the incomplete shielding of nuclei by their electrons as the vicinal action which operates.

IV. Interpretation of Experimental Facts concerning Optical Rotation

There is a great mass of experimental data concerning optical rotatory power and the many factors which can affect it. We shall now seek to interpret some of the more important of these in the light of the theories mentioned on the previous pages.

Before studying the individual factors which may operate to alter the

optical rotation of a molecule, it will be instructive to point out just where in equation 15a we might expect variable factors to occur.

$$[M] = \frac{96\pi N}{hc} \frac{n^2 + 2}{3} \sum_i \rho_i \sum_{b_i} \text{Im} \left[(a_i/\mathbf{p}/b_i) \cdot (b_i/\mathbf{m}/a_i)\right] \frac{\nu^2}{\nu_{a_i b_i}^2 - \nu^2} \quad (15a)$$

The effect of changing the frequency, ν, of the light used has already been mentioned. Relationships of the form of equation 15a are known experimentally to give the dependence of rotation on frequency with very high accuracy (25). The other factors which can alter the rotation include the refractive index, n, of the medium, the relative probabilities, ρ_i, of the different conformations, i, and the natures of the electronic states, a_i and b_i. There are thus these three ways in which the optical rotation at a given wave length can be altered, and all of the different factors which are known to be capable of varying the optical rotation must operate through them.

The sensitivity of the rotation to influences such as temperature, solvent action, and small changes in structure is annoying to one interested in forming a complete and accurate theory, but this same sensitivity makes it a potentially powerful tool in investigating the more or less minor alterations which other molecular properties (such as the refractivity) are incapable of disclosing.

A. EFFECT OF TEMPERATURE

In the liquid or solid as distinguished from the gaseous state we must include in the specifications of the conformations, i, not only the different possible orientations of the groups within the molecule, but the different possible orientations of the neighboring molecules as well, since these latter may also influence the optical rotation. Then if the conformation, i, of the entire complex has a free energy, F_i, per mole, from statistical mechanics we know that

$$\rho_i = \frac{e^{-F_i/RT}}{\sum_j e^{-F_j/RT}} \quad (38)$$

so that we obtain for the rotation of the actual mixture of conformations

$$[M] = \frac{96\pi N}{hc} \frac{n^2 + 2}{3} \frac{1}{\sum_j e^{-F_j/RT}} \sum_i \left[e^{-F_i/RT} \sum_{b_i} \frac{\nu^2 \text{Im}[(a_i/\mathbf{p}/b_i) \cdot (b_i/\mathbf{m}/a_i)]}{\nu_{b_i a_i}^2 - \nu^2} \right]$$

$$= \frac{\sum_i [M]_i e^{-F_i/RT}}{\sum_j e^{-F_j/RT}} \quad (39)$$

where $[M]_i$ is the molecular rotation corresponding to the pure complex i.

It was long considered that the unusual variations of the rotations of such substances as tartaric acid when temperature or even other factors were changed was due to the presence of several distinct molecular species in equilibrium. We see from the above and from the theoretical considerations of the earlier pages that as a result of the sensitivity of the optical rotation to molecular conformation, no very drastically different molecular species are necessary to account for these effects. A slightly different internal orientation or an altered amount of solvent effect may be sufficient to cause a marked change in the rotatory power.

As an example of how the temperature variation of optical rotation may be treated to obtain interesting information, we here consider the data of Winther (44) on the optical activity of the dimethyl, diethyl, and dipropyl esters of tartaric acid. The variation of the optical rotations of these esters with temperature may be interpreted roughly by assuming a simple reaction

B (low-temperature form) $\rightleftarrows$ A (high-temperature form)

to be the one chiefly responsible for the observed effects. The equilibrium constant for this reaction may be written:

$$\frac{b_\lambda - [\alpha]_\lambda^T}{[\alpha]_\lambda^T - a_\lambda} = e^{\frac{\Delta S}{R}} e^{-\frac{\Delta H}{RT}} \tag{40}$$

where a_λ is the rotation of pure liquid A, b_λ is that of pure B, and $[\alpha]_\lambda^T$ is the observed rotation of the equilibrium mixture at a temperature T; all rotations are measured using light of wave length λ.

Values of the unknown constants of equation 40 were obtained which would give a fair agreement between observed and calculated rotations at various wave lengths. The values of ΔH are probably good to about ± 400 cal.; ΔS may be in error by ± 2 units. The values of a_λ and especially of b_λ are strongly dependent on the values of ΔS and ΔH taken, but the relative values at different wave lengths for a particular choice of ΔS and ΔH are probably more reliable. The results are summarized in table 2. Since ΔS and ΔH were chosen to give a good fit at one particular wave length and the same values were used at other wave lengths with values of a_λ and b_λ which give good agreement at but two temperatures, the agreement is better at some points than at others. Slightly different values of the constants would have given more even agreement, but the values used are sufficiently accurate for our purpose, which is that of obtaining an estimate of the heat and entropy changes involved.

The close similarity between the constants for the three esters leaves little doubt as to the essential correctness of the above interpretation of Winther's results. The large change in entropy between the two forms

ESTER	ΔS	ΔH	SPECIFIC ROTATIONS OF FORMS A AND B			
	E. U.	*calories*				
Dimethyl tartrate	10.6	2280	λ = 5890 Å.	λ = 4703 Å.	λ = 4445 Å.	
			$a = +10.2°$ $b = -30°$	$a = +10.94°$ $b = -61.4°$	$a = +11.1°$ $b = -81°$	
Diethyl tartrate	9.8	2530	λ = 5890 Å.	λ = 5335 Å.	λ = 4655 Å.	λ = 4435 Å
			$a = +18.9°$ $b = -13.6°$	$a = +22.1°$ $b = -19.1°$	$a = +26.5°$ $b = -34.9°$	$a = +27.4$ $b = -41.8$
Dipropyl tartrate	8.4	2300	λ = 5890 Å.	λ = 4445 Å.		
			$a = +24.01°$ $b = -4.60°$	$a = +37.9°$ $b = -22.68°$		

Dimethyl tartrate

λ = 5890 Å.			λ = 4703 Å.			λ = 4445 Å.		
T	$[\alpha]_{calcd.}$	$[\alpha]_{obsd.}$	T	$[\alpha]_{calcd.}$	$[\alpha]_{obsd.}$	T	$[\alpha]_{calcd.}$	$[\alpha]_{obsd.}$
°*K.*			°*K.*			°*K.*		
325.3	4.22	4.22	323.8	0.02	0.02	324.9	−2.64	−2.62
334.2	4.69	4.68	335.1	1.11	1.11	335.1	−1.43	−1.14
344.9	5.17	5.16	344.3	1.85	1.74	344.9	−0.06	−0.50
354.8	5.57	5.56	355.6	2.62	2.62	356.0	+0.58	+0.53
365.7	5.96	5.96	364.3	3.22	3.23	364.8	+1.27	+1.31

Diethyl tartrate

λ = 5890 Å.			λ = 5335 Å.			λ = 4655 Å.			λ = 4435 Å.		
T	$[\alpha]_{calcd.}$	$[\alpha]_{obsd.}$	T	$[\alpha]_{calcd.}$	$[\alpha]_{obsd.}$	T	$[\alpha]_{calcd.}$	$[\alpha]_{obsd.}$	T	$[\alpha]_{calcd.}$	$[\alpha]_{obsd.}$
°*K.*			°*K.*			°*K.*			°*K.*		
292.1	7.30	7.31	292.1	7.16	7.16	292.2	4.57	4.57	292.2	2.70	2.7
301.2	8.27	8.26	303.9	8.80	8.71	302.7	6.64	6.61	303.2	5.12	5.0
313.1	9.37	9.38	311.2	9.68	9.59	312.5	8.37	8.21	314.1	7.27	7.2
325.0	10.33	10.35	321.4	10.73	10.69	322.8	9.98	9.89	324.7	9.14	8.8
336.4	11.13	11.15	334.5	11.99	11.93	336.3	11.75	11.64	335.6	10.73	10.6
344.3	11.63	11.65	343.6	12.75	12.76	345.2	12.84	12.78	344.4	11.93	11.8
352.9	12.13	12.15	354.8	13.56	13.57	354.6	13.86	13.86	354.3	13.18	13.1

Dipropyl tartrate

λ = 5890 Å.			λ = 4445 Å.		
T	$[\alpha]_{calcd.}$	$[\alpha]_{obsd.}$	T	$[\alpha]_{calcd.}$	$[\alpha]_{obsd.}$
°*K.*			°*K.*		
289.1	11.36	11.39	289.1	11.14	11.14
301.1	12.48	12.58	302.6	13.73	13.65
312.1	13.40	13.41	312.4	15.45	15.56
321.8	14.15	14.22	322.4	17.04	17.18
335.4	15.05	15.10	335.4	18.90	18.99
343.9	15.56	15.57	346.1	20.25	20.26

leads us to suspect that they differ in amounts of solvation or, more probably, of free rotation, rather than merely in static configuration.

Lucas (27) has pointed out that a maximum in the temperature *versus* rotation curve indicates the presence of at least three distinct substances in equilibrium. Such maxima are known to occur in the curves for the tartaric acid esters, so Lucas suggests that there must be three forms of these esters. On examining the actual data (32), however, it is found that the "maxima" are readily accounted for by the indirect effect of the density acting through the refractive index (see page 385), and that when account is taken of this, the maxima usually disappear.

B. EFFECTS OF ROTATION ABOUT BONDS AND SYMMETRY OF GROUPS

In discussing the modern theories of optical activity it was noticed that the theories of Kuhn and Kirkwood, as well as the one-electron theory when only first-order effects are considered, all assume that vicinal effects are additive; that is, the addition of a third group does not affect the interactions already existing between two other groups. We shall now examine some of the consequences of this assumption.

First, consider the interactions between atoms or groups A and B both attached to the same atom C, where the lines AC and BC are axes of symmetry of groups A and C, respectively $\left(\text{e.g., } \mathrm{C}\!\begin{smallmatrix}\diagup\mathrm{Cl}\\ \diagdown\mathrm{Br}\end{smallmatrix}\right)$. Since such an arrangement of groups has a plane of symmetry, no interaction between A and B can alone give rise to optical activity.

Next consider four atoms arranged as shown in figure 8. (A plane passes through B, C, and D.) The interaction of AB with D when A is above the plane (position A′) is equal and opposite in sign to that when A is below the plane (in position A″, which is the reflection of A′ in the plane). Therefore, if positions A′ and A″ are equally probable, these interactions will cancel off. And if all other possible positions of A above the plane are matched by equally probable positions below the plane, the interactions of AB and D will not influence the optical activity.

The argument can be extended to groups of any size, as long as these groups can be split up into interacting units having planes of symmetry. It also holds true if there is restricted rotation in which the positions of minimum potential energy are equally probable on either side of a plane which passes through the asymmetric center and the group with which the given group is interacting, this being the situation in the figure accompanying the last paragraph.

The following are some of the consequences of the above: (*a*) Factors which decrease the freedom of rotation of groups about bonds will cause

an increase in the first-order contribution to the optical activity. (*b*) Insofar as an increase in temperature causes an increase in the amount of rotation of groups about bonds in such molecules as *sec*-butyl chloride, their rotations should tend toward a value corresponding to the second-order contributions as the temperature rises. (*c*) Any optical rotation possessed by such compounds as CHClBrI must be due to second-order effects.

Now we have seen (page 361) that second-order effects are in general smaller than first-order effects, especially as the groups which interact to give rise to optical activity become more and more widely separated. It will, then, be interesting to see if the conclusions which we can draw from this with *a*, *b*, and *c* above, agree with those found experimentally.

(*a*) A powerful factor in reducing the possibility of rotation about single bonds is the formation of ring compounds from open chains, and it should, from the above, be accompanied by an increase in optical rotation. The

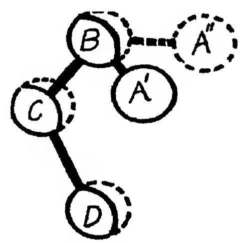


FIG. 8

fact that ring formation greatly increases the optical rotation has long been known experimentally (42).

In table 3 are given the rotations of compounds chosen at random which are similar in all respects save that some have open chains while others have closed rings. The influence of ring formation is found generally to be as predicted. Furthermore, the effect is the most pronounced in those compounds in which there is the most reason to believe that there is the greatest freedom of rotation in the open-chain form. Thus the introduction of the bulky acetyl in place of a hydrogen on a hydroxyl raises the rotation somewhat. (Even here, however, some of the increase is due to the fact that the acetyl has an absorption band nearer to the visible than does hydroxyl or the hydrocarbon residue.) A methyl group has a similar but less marked effect. The very low rotations of the polyhydroxy alcohols make it apparent that here there must be either very nearly free rotation or near-equivalence of the three equilibrium positions about each C—C and C—O bond.

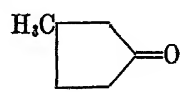
TABLE 3

Comparison of rotations of open-chain compounds and ring compounds of similar constitution

OPEN-CHAINS	$[M]_D$	RINGS		$[M]_D$
Arabite, OHOHH CH₂OH—C—C—C—CH₂OH H H OH	8.4°	Ribose, OHOHOH CH₂—C—C—C—CHOH H H H O		32.3°
Talite, H OHOHOH CH₂OH—C—C—C—C—CH₂OH OHH H H	5.5°	Arabinose, OHOHH CH₂—C—C—C—CHOH H H OH O	α-form.......... β-form..........	263° 75°
Mannite, H H OHOH CH₂OH—C—C—C—C—CH₂OH OHOHH H	0.9°	Xylose, H OHH CH₂—C—C—C—CHOH OHH OH O	α-form..........	138°
Idite, OHH OHH CH₂OH—C—C—C—C—CH₂OH H OHH OH	6.4°	Mannose, H H OHOH CH₂—C—C—C—C—CHOH OH OHH H O	α-form.......... β-form..........	54° 24.5°
Sorbite, H H OHH CH₂OH—C—C—C—C—CH₂OH OHOHH OH	3.6°	Galactose, H OHOHH CH₂—C—C—C—C—CHOH OH H H OH O	α-form.......... β-form..........	260° 93.5°
		Glucose, H H OHH CH₂—C—C—C—C—CHOH OH OHH OH O	α-form.......... β-form..........	203° 34°

			α-form	β-form
Mannose dimethylacetal	1.3°	Methyl mannopyranoside	159°	
(pentaacetate	69°)	(tetraacetate		170°)
Galactose dimethylacetal	34°	Methyl galactopyranoside	307°	62°
(pentaacetate	?)	(tetraacetate	474°	65°)
Glucose dimethylacetal	34°	Methyl glucopyranoside	345°	1.0°
(pentaacetate	52°)	(tetraacetate	478°	50°

Arabonic acid	$[\alpha]_D = 18° \rightarrow \gamma$-Lactone, $[\alpha]_D = 74°$
2,3,4-Trimethylarabonic acid	$[\alpha]_D = 16° \rightarrow \gamma$-Lactone, $[\alpha]_D = 178°$
2,3,5-Trimethylarabonic acid	$[\alpha]_D = 2° \rightarrow \gamma$-Lactone, $[\alpha]_D = 44°$
2,3,4-Trimethyllyxonic acid	$[\alpha]_D = 1.3° \rightarrow \gamma$-Lactone, $[\alpha]_D = 35°$
2,3,5-Trimethyllyxonic acid	$[\alpha]_D = 21° \rightarrow \gamma$-Lactone, $[\alpha]_D = 82°$
3,4-Dimethylrhamnonic acid	$[\alpha]_D = 16° \rightarrow \gamma$-Lactone, $[\alpha]_D = 153°$
Rhamnonic acid	$[\alpha]_D = 7.7° \rightarrow \gamma$-Lactone, $[\alpha]_D = 40° \rightarrow \delta$-Lactone, $[\alpha]_D = 98°$
Gluconic acid	$[\alpha]_D = 7° \rightarrow \gamma$-Lactone, $[\alpha]_D = 68° \rightarrow \delta$-Lactone, $[\alpha]_D = 62°$
Mannonic acid	$[\alpha]_D = 10° \rightarrow \gamma$-Lactone, $[\alpha]_D = 52° \rightarrow \delta$-Lactone, $[\alpha]_D = 112°$
Galactonic acid	$[\alpha]_D = 13° \rightarrow \gamma$-Lactone, $[\alpha]_D = 77°$
2,3,5,6-Tetramethylgluconic acid	$[\alpha]_D = 34° \rightarrow \gamma$-Lactone, $[\alpha]_D = 72°$
2,3,4,6-Tetramethylgluconic acid	$[\alpha]_D = 22° \rightarrow \gamma$-Lactone, $[\alpha]_D = 101°$

$CH_3CH_2CH(CH_3)CH_2CHO$	$[M]_D = 8.7°$		$[M]_D = 120°$

If we say that, on the average, ring formation causes a tenfold increase in optical rotation, and that this represents the difference in order of magnitude between first- and second-order effects, then this corresponds to values of the coefficients c in equation 28 of about 0.1,—a reasonable value such as has been found in actual calculations.

Another factor besides ring formation which acts to hinder the movement of one part of a molecule relative to another is the formation of the crystalline state. Therefore we expect that solidification will cause an increase in optical rotation. Thus for quartz, sodium perchlorate, etc., the optical rotation goes from zero for the melt to exceedingly high values for the crystal. For tartaric acid, the amorphous acid at room temperature has

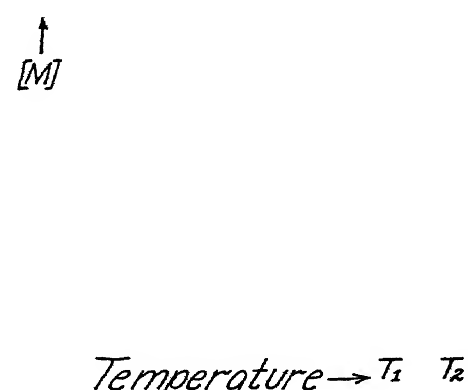


FIG. 9. Effect of temperature on rotatory power

a specific rotation of $+0.76°$ at $\lambda = 5920$ Å. and $+15.9°$ at 180°C. for $\lambda = 5880$ Å., while the crystalline acid at room temperature has a specific rotation of $-636°$ for $\lambda = 5780$ Å.

(*b*) If we restrict ourselves to substances whose second-order contribution to the optical rotation should be small (as in $CH_3CH_2CHClCH_3$, where the groups responsible for the asymmetry,—namely, the CH_3 of the ethyl, the H of the methyl, and the H and Cl directly attached to the asymmetric center,—are considerably farther apart than the corresponding ones in, say, tartaric acid) then we must conclude that an increase in temperature will generally cause a considerable decrease in the numerical value of the rotation. Since the sign of the second-order contribution to the optical rotation need not be the same as that of the first-order contribu-

tion, this is not to say that over a small range of temperatures an increase in the rotatory power with increasing temperature will not be observed. Thus, in figure 9, although there is a change in rotation from a large to a small value, measurements in the interval T_1 to T_2 would show an increase in rotatory power with increasing temperature.

Now it has been noted in the past that the optical rotations of simple substances generally seem to decrease with increasing temperatures. Guye and Aston (14) found that for every one of thirty substances which meet the requirement given above, a temperature increase lowers the rotation. Several other substances noted by Walden (43) show the same effect. The magnitude of the decrease with temperature is, however, usually rather small, so that one wonders whether in many of the cases cited, at least, the decrease is not caused largely by a decrease in the refractive index consequent on the lower density at higher temperatures (see later). If this is the case, we are forced to the interesting conclusion that in these substances the molecules already possess freedom of rotation and that an adequate test of the principle is only possible with those substances having groups sufficiently bulky so that freedom of rotation only occurs at higher temperatures. A study of substances of this type is now being made, and preliminary findings tend to confirm expectations from the above.

(*c*) Recently, Berry and Sturtevant (2) have prepared $CH_3CHBrCN$ in optically active form. This substance was found to have a molecular rotation in sodium D light of at least 20.5°, and since the bonds joining the groups to the center of asymmetry all lie along axes of symmetry of the respective groups, this must be entirely a result of contributions of second order and higher. The proximity of the groups, along with the nearness of the absorption bands of Br and CN to the visible, makes this value quite reasonable. (In the polyhydroxy alcohols and the substances studied by Guye and Aston the value seems generally to be of the order of 10° or less.) It is worth noting that the value 20.5° is still very much smaller than the rotation of 120° by the even simpler methylcyclopentanone; in this compound there is a first-order contribution to the optical rotation (see later), and the comparison is the more striking when it is learned that this large rotation arises from an interaction over the very considerable distance of 4.5 Å., whereas in $CH_3CHBrCN$ the distances are of the order of 2.5 to 3 Å.

C. SOLVENT EFFECTS

In this discussion the solvent is considered as being made up of all those molecules which surround a given optically active molecule. These neighboring molecules may or may not be of the same species as the active molecule. There is obviously no distinction in principle between

solvent effects of, for example, water on *sec*-butyl alcohol and the effects of one *sec*-butyl alcohol on another. Thus solvent effects will occur in all states except the dilute vapor.

It will be convenient to divide the effects which may act here into two types according to origin; the first will be the effect of refractive index, and the second the effects of alterations of β, which are more deep-seated, hence more interesting.

1. Effect of refractive index

The refractive index, n, enters into the expression for optical activity as a factor $(n^2 + 2)$, or better as $(n^2 + 2)/3$, since this reduces to unity for the vapor. Therefore, if β in equation 15 is a constant, the quantity $[M]/(n^2 + 2)$ should be a constant, in going from one solvent to another. Wolf and Volkmann (45) conclude that where non-polar solutes are involved, this quantity actually is a constant. Beckmann and Cohen (1) conclude the same and suggest that in investigating solvent effects on optical activity the variation of

$$\Omega = \frac{[\alpha]}{n^2 + 2} \tag{41}$$

should be studied. (For the reason mentioned above, it might be better to use three times this quantity, it having a value more nearly related to that of the conventional $[\alpha]$.) They call Ω the *rotivity* in analogy to the relation between the refractive index and the refractivity. The reasoning behind the use of Ω instead of $[\alpha]$ is sound, since Ω is proportional to the more fundamental molecular quantity β and its variation is of greater significance than the variation of $[\alpha]$.

The factor $(n^2 + 2)$ in the optical rotation, as in the theory of dipole moments, corrects the local field around a molecule for the polarization of neighboring molecules.

That β itself may in some way depend upon the refractive index is indicated by the results of Rule and Chambers (37) on the effect of refractive index on the rotation of the saturated hydrocarbon pinane in various solutions. In table 4 and figure 10 are given the dependence of $[M]_D^{20}$ of pinane solutions on n_D^{20}, the refractive index of the solution, on $(n^2 + 2)/3$, and on $[(n^2 + 2)/3]^2$. Although all of the variable factors have not been accounted for, the most important one in this case is clearly the refractive index. The major portion of this dependence appears to be eliminated when account is taken of the factors $(n^2 + 2)/3$ or $[(n^2 + 2)/3]^2$. Both of these factors give some residual dependence on the refractive index, but when account is taken of the fact that p-fold division by $(n^2 + 2)/3$ tends of itself to give a smaller slope to any plot of

TABLE 4

Refractive indices and rotations of solutions of pinane (reference 37)

SOLVENT	n_D^{20} OF SOLUTION	$[M]_D^{20}$	$[M]_D^{20} \div \left(\frac{n^2+2}{3}\right)$	$[M]_D^{20} \div \left(\frac{n^2+2}{3}\right)^2$
Methyl cyanide	1.3441	25.8	20.4	16.1
Methyl alcohol	1.3266	26.4	21.1	16.8
Acetic acid	1.3718	26.9	20.8	16.1
Nitromethane	1.3864	27.3	20.9	16.0
Pentane	1.3640	27.4	21.3	16.6
Acetaldehyde	1.3316	27.7	22.0	17.5
Hexane	1.3835	27.8	21.3	16.3
Acetone	1.3653	28.9	22.5	17.5
Methylene chloride	1.4245	29.0	21.6	16.1
Chloroform	1.4489	30.2	22.1	16.2
Phenyl cyanide	1.5255	30.8	21.3	14.7
Methylene bromide	1.5385	31.4	22.2	15.2
(Homogeneous)	1.4630	31.5	22.8	16.5
Carbon tetrachloride	1.4616	31.6	22.9	16.6
Pyridine	1.5088	32.3	22.6	15.9
Mesitylene	1.4944	32.5	23.1	16.4
Nitrobenzene	1.5468	32.5	22.2	15.2
Methyl iodide	1.5291	32.6	22.5	15.6
Benzene	1.4992	32.9	23.0	16.2
Acetophenone	1.5310	33.2	23.0	15.9
Benzaldehyde	1.5431	33.2	22.8	15.6
Toluene	1.4930	33.7	24.0	17.0
Anisole	1.5141	34.5	24.1	16.8
Chlorobenzene	1.5229	34.7	24.1	16.7
o-Dichlorbenzene	1.5466	34.7	23.7	16.2
α-Chloronaphthalene	1.6332	34.9	22.4	14.4
Bromobenzene	1.5580	35.1	23.8	16.1
α-Bromonaphthalene	1.6558	35.1	22.2	14.1
Ethylene dibromide	1.6303	35.3	22.7	14.6
Iodobenzene	1.6162	35.6	23.2	15.1
α-Methoxynaphthalene	1.6201	35.6	23.1	15.0
Aniline	1.5842	36.4	24.2	16.1
Methylene iodide	1.7341	36.6	22.0	13.2
Quinoline	1.6246	38.6	25.0	16.2
Carbon bisulfide	1.6246	38.7	25.0	16.2
α-Iodonaphthalene	1.7054	39.7	24.8	15.2

$[M] \div \left(\frac{n^2+2}{3}\right)^P$ *versus* n, it is found that the rotation[4] is more nearly proportional to $(n^2 + 2)$ than to $(n^2 + 2)^2$.

[4] Rule and Chambers find that for solvents with zero dipole moment the rotation is more nearly proportional to $(n^2 + 2)^2$ than to $(n^2 + 2)$. The dipole moment, however, does not appear to be related to those factors which, besides the refractive

Often data concerning refractive indices are not available, so that the rotivity cannot be found. In case the dependence of rotation on tempera-

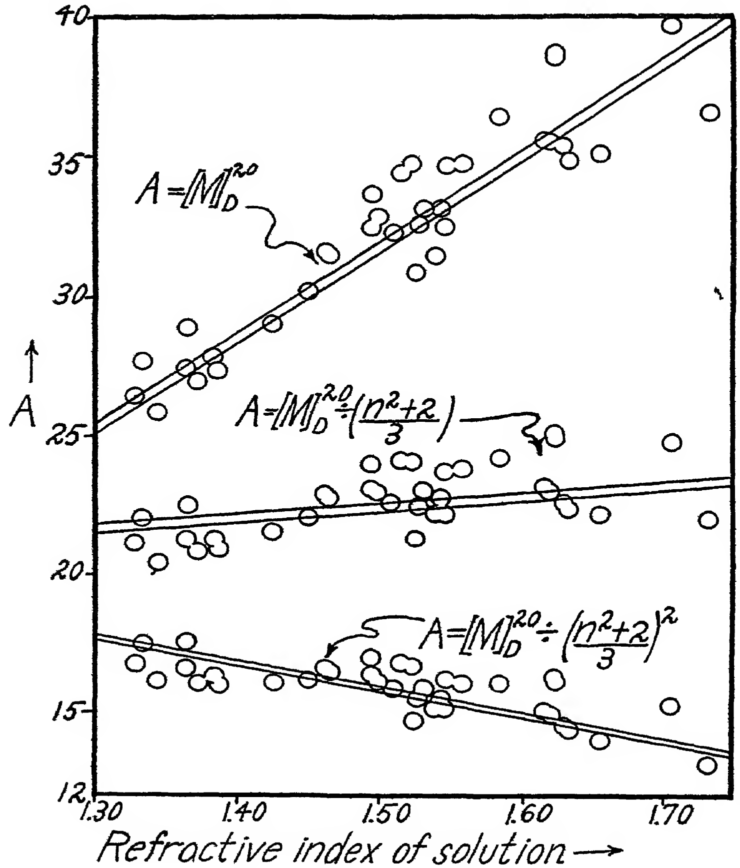


FIG. 10. Dependence of rotation of *d*-pinane on refractive index of its solutions. Distance between parallel straight lines indicates magnitude of probable experimental error.

ture is being investigated, however, the following method may prove satisfactory: We know that the Lorenz–Lorentz equation,

$$\frac{n^2 - 1}{n^2 + 2}\frac{M}{d} = P$$

index, can alter the rotation of pinane, so that the basis of the criterion upon which they base their conclusion is not valid. A dependence of V in equations 23, 32, and 33 on the dielectric constant, hence on the refractive index, might, however, cause a trend such as is observed in the rotivity *versus* refractive index plot.

where P is a constant, is accurately obeyed by most substances under varying conditions. For solutions, M and P may be calculated in the same way as any colligative property, or P/M may be found if the refractive index and density are known at any temperature. Now, from the Lorenz–Lorentz equation it is found that

$$n^2 + 1 \qquad \overline{1 - (P/M)d} \tag{42}$$

so that once P/M is known, the rotivity of a given solution may be found at any temperature from a knowledge of the density variation alone. An example is given in table 5. Here a maximum in the rotation as the temperature increases does not appear in the rotivity.

TABLE 5

Effect of refractive index changes on rotation of ethyl tartrate at different temperatures

$P = 45.70$ from atomic refractions; $M = 206$; $P/M = 0.222$

t	d	$[M]^t_{5761}$	$(1 - 0.222d)[M]^t_{5761}$
°C.			
−22.0	1.2472	3.84	2.78
0.0	1.2254	11.03	8.03
15.5	1.2097	15.09	11.04
59.5	1.1656	23.77	17.62
114.9	1.1095	29.60	22.31
190.0	1.0345	31.99	24.64
223.5	1.0003	31.75	24.70

Data from reference 32.

2. *Solvent effects which act through β*

Changes in the rotivity must all result from alterations in the states a_i and b_i of the molecule. These alterations must in turn arise from compound formation of one sort or another, that is, there must be formed some kind of bonds, be they loose dipole–dipole bonds, or strong "chemical" bonds, between the solvent molecules and the active molecule. We shall here distinguish between the following mechanisms by which such compound formation alters the rotation: (*a*) action of a group or groups in the solvent molecule as chromophores; (*b*) direct vicinal action of the solvent molecule on the chromophores of the active molecule; and (*c*) distortion of the molecular framework of the active molecule by the solvent and consequent effect on the optical rotation.

(*a*) *Action of groups of the solvent molecule as chromophores* will certainly always occur to a considerable extent when a covalent bond between the solvent and the solute is formed. An example might be found in the fact

that on addition of chromium ions to tartaric acid solutions, the absorption bands of chromium show anomalous dispersion of the optical rotation due to the formation of a chromi-tartrate complex. But of the very great and irregular effects of solvents which undergo a chemical reaction with the optically active molecule we need only remark that they are entirely understandable as compared with the more obscure and smaller variations caused by solvents which certainly undergo no ordinary chemical reaction with optically active molecules upon which they can act. It is therefore the latter which will interest us here.

Unless there is a very strong bond between the solvent molecule and the optically active molecule, it is questionable if the solvent will act appreciably as a chromophore. First of all, the many possible orientations of nearly the same energy which the solvent molecules could take on about an optically active molecule with which they form no complex of definite structure would tend to cause the contributions from this source to cancel out. Secondly, vicinal actions capable of causing the solvent to act as a chromophore fall off rapidly with distance, and two neighboring molecules in a solution are usually relatively farther apart than those groups in the molecule whose interactions determine the optical rotation.

Experimental evidence for the deduction that a solvent may have an effect on the optical rotation without itself acting as a chromophore is found in the work of Lifschitz (24). Lifschitz finds that several metal ions which have an appreciable effect on the rotation of oxymethylenecamphor and which have absorption lines in the visible show no anomalous dispersion in the neighborhood of these lines.

(b) and (c). Alteration of β both by direct vicinal action of the solvent molecule on the chromophores of the active molecule and by distortion of the molecular framework of the active molecule may occur and both will be considered together in the following treatment. Here the theory of Beckmann and Cohen (1) finds its application.

These writers begin by assuming that the rotivity of a molecule is linearly proportional to the electrostatic field acting on the active molecule. That is,

$$\Omega = \Omega_0 + \Omega' F \tag{43}$$

where Ω' is the change in rotivity per unit field acting along a certain chosen direction. We shall see later that we need only be concerned with fields which act along the direction of the resultant dipole of the optically active molecule, so that Ω' will be defined here with respect to that direction.

It is instructive to ask if the theories of optical activity which we have outlined lead us to expect such a linear relationship. In the first place,

we note again that since the development is made in terms of the rotivity, we are directly concerned with effects involving β, the molecular rotatory parameter. (Beckmann and Cohen, however, regarded the rotivity as a measure of the third-rank Darwin scattering factor, which they treat as a molecular constant.) We must therefore ask how β will be expected to depend on the field. The following mechanisms suggest themselves: first, the applied field might distort the structural framework of the molecule. Such distortions would affect the optical rotation, since this is so sensitive to the positions of atoms and groups with respect to the various chromophoric groups of the molecule. If the force resisting distortion obeys Hooke's law and if the force acting to cause distortion is proportional to the field (both assumptions are reasonable), there should be a linear relationship between the field and the amount of displacement of the groups of the molecule, even for quite large displacements. Now over a sufficiently small range of displacements it is possible to say that the effect on the optical rotation will be proportional to the amount of displacement. Therefore we have that the change in rotivity will be proportional to the field.

A second mechanism whereby a field might influence the optical rotation is illustrated by the following example: Suppose that there are two positions, A and B, in which a group is predominantly found with respect to rotation about a bond. Then the optical rotivity will be given by $\Omega = n_A\Omega_A + n_B\Omega_B$ where n_A and n_B are the fractions of molecules having the group in question in positions A and B, respectively, while Ω_A and Ω_B are the rotivities corresponding to each position. Now, in a field F position A may become more stable by an energy d_AF, while position B may become less stable by an energy d_BF, due to interaction of dipoles or polarizability ellipsoids with the field. n_A will then be increased by an amount proportional to $e^{\mathrm{d}_A F/kT}$, while n_B will be decreased by an amount proportional to $e^{\mathrm{d}_B F/kT}$. If the energies d_AF and d_BF are small relative to kT, the exponentials may be expressed as $(1 + \mathrm{d}_AF/kT)$ and $(1 + \mathrm{d}_BF/kT)$. Introducing these into our expression for Ω in terms of n_A and n_B, we see that the rotivity will depend linearly upon the field.

A third mechanism, distinct from the above, suggests itself when we inspect the experimental data concerning the relative orders of magnitude of the effects of solvents on rigid and non-rigid molecules. If the mechanisms proposed above were the only ones operating, we should expect the rotations of such molecules as camphor and fenchone, which possess cross-braced, rigid structures, to show very much less susceptibility to solvents than non-rigid substances such as menthylmethyl naphthoate. This is often found not to be the case; some rigid molecules show large solvent influences, while some non-rigid molecules show small ones. This would

indicate that the solvent field is exerting some direct vicinal action on the chromophores of the active molecule. Such actions would be expected from the one-electron theory and have been shown in section III C to have a linear effect on the rotation to a first and probably good approximation. It is to be noted that arguments of symmetry similar to those of IV B lead to the expectation that such vicinal actions will be greatest when they do not lie along the axes of the more strongly chromophoric groups. That is, the direction of the resultant dipole moment of the active molecule must lie in a different direction from those of the axes of the important chromophoric groups. It is of interest that those molecules whose rotations show the greatest susceptibility to solvent action generally appear to conform to this requirement.

In order to calculate the rotivity for an actual system of molecules, using equation 43, it is necessary to weight the rotivity corresponding to each configuration of the active molecule and its neighbors by the probability of occurrence of the configuration. For the term in the rotivity proportional to the field, the same result is obtained by simply averaging the field over all configurations and introducing the resultant value of the field directly into equation 43. This is accomplished as follows: If the active molecule has a dipole moment μ_α, it will induce a moment in the surrounding molecules and it will also interact with their permanent moments. Each configuration of the system gives an electric field acting on the optically active molecule due to this polarization of the surrounding molecules, and each configuration has a probability of occurring which is proportional to $e^{-\epsilon/kT}$, where ϵ is the energy of the configuration. An average field is obtainable by the usual statistical mechanical methods. Assuming the molecules to be hard spheres with dipoles at their centers, and assuming the interaction energies to be small relative to kT, the average field will be in the direction of the dipole of the active molecule (whence the form of our definition of Ω' at the start) and has the value,

$$F_{av.} = \frac{2\mu_\alpha}{d^3_{\alpha\alpha}} \frac{n_\alpha}{N} P_\alpha + \frac{2\mu_\alpha}{d^3_{\alpha\beta}} \frac{n_\beta}{N} P_\beta + \cdots \tag{44}$$

where n_α = number of optically active molecules per unit volume,
n_β = number of solvent molecules of type β per unit volume,
N = Avogadro's number,
$d_{\alpha\alpha}$ = distance of closest approach of two active molecules, taken as hard spheres,
$d_{\alpha\beta}$ = distance of closest approach of an active molecule and a solvent molecule of type β,
$P_\alpha = \frac{4\pi N}{3}\left[A + \frac{\mu_\alpha^2}{3kT}\right]$, where A = mean polarizability of an active molecule,

$P_\beta = \frac{4\pi N}{3}\left[B + \frac{\mu_\beta^2}{3kT}\right]$, where B = mean polarizability of a solvent molecule of type β, and

μ_α, μ_β = dipole moments of active and solvent molecules.

It is to be noted that A and B are independent of the concentrations of the molecules in the solution to a very good approximation (38). Owing to the association of molecules, however, μ_α and μ_β may depend upon the concentration, but as long as the result of this association is the same for the microscopic field about an active molecule as for the microscopic field used in measuring dielectric constants, P_α and P_β may be found from the relation

$$\frac{\epsilon - 1}{\epsilon + 2} = \frac{1}{N}(n_\alpha P_\alpha + n_\beta P_\beta + \cdots) \tag{45}$$

where ϵ is the dielectric constant of the solution.

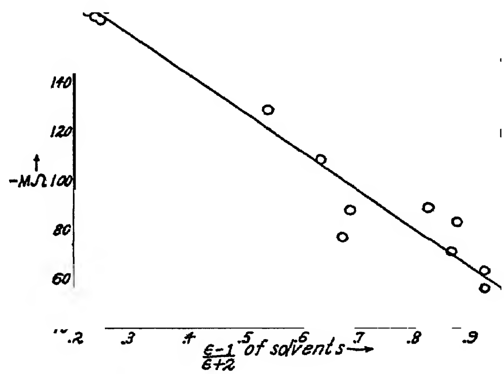


FIG. 11. Plot of Ω against $(\epsilon - 1)/(\epsilon + 2)$ for l-menthylmethyl naphthoate in aliphatic solvents

Substituting equations 44 and 45 in equation 43 and taking dilute solutions of the optically active molecule of a definite concentration, we find

$$\Omega = \Omega_0 + \Omega' \cdot \frac{2\mu_\alpha}{d^3} \frac{\epsilon - 1}{\epsilon + 2} \tag{46}$$

where d is approximately constant for solvents of similar structures. Taking ϵ for the pure solvent, Beckmann and Cohen find that Ω plotted against $\frac{\epsilon - 1}{\epsilon + 2}$ gives a fairly good straight line (see figures 11, 12, and 13) for several substances investigated by Rule.

The deviations from linearity and the fact that a distinction must be

made between aromatic and aliphatic solvents show that other effects must be active but that the main effect is probably due to the action of dipoles in the manner pictured above. Furthermore, the independence of $\frac{\epsilon - 1}{\epsilon + 2}$ and the rotivity of pinane, whose dipole moment is zero, follows from equation 46.

The connection between the dipole moment of solvent molecules and their effect on the rotation had been demonstrated previously by Rule and his coworkers (35). This is now understandable, since the major contribution to the polarization of polar substances occurs through their dipole moments.

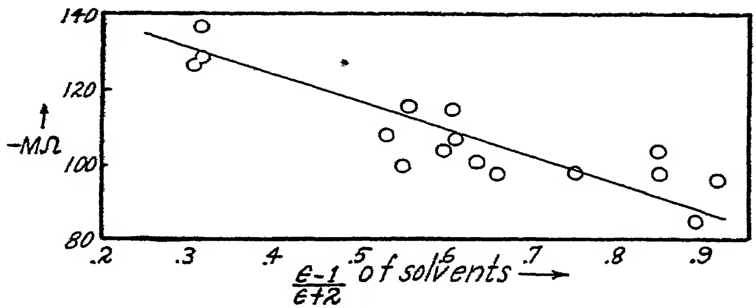


FIG. 12. Plot of Ω against $(\epsilon - 1)/(\epsilon + 2)$ for *l*-menthylmethyl naphthoate in aromatic solvents

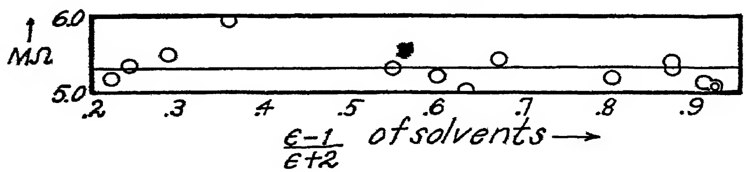


FIG. 13. Plot of Ω against $(\epsilon - 1)/(\epsilon + 2)$ for *d*-pinane

Beckmann and Cohen performed experiments in which they measured the optical rotations of active substances having a fixed concentration in mixtures of polar and non-polar solvents of different concentrations. They verified that there is a linear relation between the change in rotivity per unit amount of polar solvent and the polarization of the polar solvent at a given concentration. That is, as can readily be shown from their theory,

$$\frac{\Omega - \Omega_0}{n} = G + KP \tag{47}$$

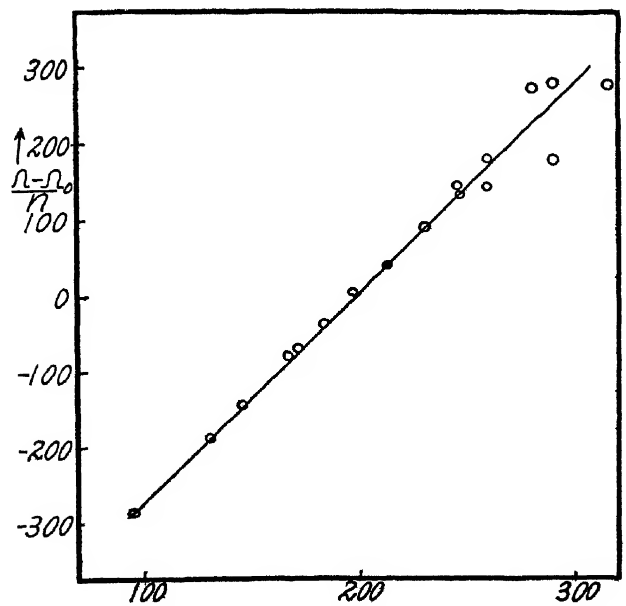
where n = concentration of polar substance in moles per cubic centimeter,

Ω = rotivity of a fixed concentration of active substance in the mixed solvent,

Ω_0 = rotivity of active substance at the same concentration in pure non-polar solvent,

G and K are constants depending on the nature of the polar substance, among other things, and

P = molecular polarization of polar substance as found from equation 45 for each value of n.



Polarization of nitrobenzene in solution (in cubic centimeters)

FIG. 14

G measures, at least in part, the difference in the rotivity when polar molecules replace non-polar ones owing to other than polarization forces. The linearity of the relation, equation 47, does not, however, prove that this difference is really independent of the polarization, since any linear dependence is included in K. A selective crowding out of non-polar molecules by polar ones in the neighborhood of the active molecule due to differences in polarization might be expected to introduce such a linear dependence on P.

The excellence of the relationship in equation 47 is shown by the curve in figure 14 for diethyl diacetyl-d-tartrate in mixtures of benzene and nitrobenzene, varying from 2 g. of nitrobenzene per 50 cc. of solution to 100 per cent nitrobenzene as solvent, with 5 g. of tartrate per 50 cc.

throughout. Neither P nor Ω varies linearly with the concentration of nitrobenzene, yet the two are themselves related linearly. For further deductions from the theory, the reader is referred to the original articles (1).

D. EFFECTS OF POLAR GROUPS WITHIN THE ACTIVE MOLECULE

We have just seen that a field applied along a certain direction of an optically active molecule causes a linear change in the rotivity. If such a field is produced directly by introducing a dipole into the molecule itself, a similar result both as to origin and magnitude is to be anticipated. Rule and Betti have indeed sought to establish a relationship between the polarity of a group substituted at some point in an optically active molecule and the optical rotation of the molecule.

Unfortunately there are a number of factors which complicate the problem. In the first place, the various groups which are introduced may themselves act as chromophores and the resultant contribution to the optical rotation may bear no relation to the polarity of the group. Secondly, no account is usually taken of the "solvent effect" of the dipoles in one active molecule on the rotation of another molecule of the same species. This action is the same as that discussed in the previous section and is certain to occur unless rotations are measured in dilute solution. It would lead to an apparent relationship between the polarity of the group and the optical rotation which is actually no different from the question of solvent action discussed before. Thirdly, even if the above factors were eliminated, although a definite relationship between the optical rotation and the magnitude of the dipoles in a molecule is to be expected from the one-electron theory, the directness of this relationship would be reduced by the fact that not only the magnitude of the moment but its orientation in space as well must be considered.

In the light of these considerations, then, the results obtained by Rule (36) and Betti (3), some of which are given in table 6, are understandable both in the regularities and in the discrepancies which they exhibit. Since Betti's measurements were made in benzene solution, the "solvent effects" of the molecules on one another are reduced and more regularity is exhibited than in the results of Rule, whose measurements were made using the undiluted material. Furthermore, the strikingly large changes in rotation which are obtained by Betti indicate that something more than simple solvent action must be acting here.

The key to the understanding of the unmistakable relationship found by Betti between the rotation of $RCH{=}NCH(C_6H_5)C_{10}H_6OH$ and the strength of the acid RCOOH is probably to be found in the work of Bjerrum, and Kirkwood, Westheimer, and Shookhoff (4, 16) on the effects of electrostatic fields resulting from dipoles and charges in an acid on the

strength of the acid. These workers are able to account rather well for the values of the dissociation constants of acids in which there are dipoles and free charges by assuming that these, through their electrostatic fields,

TABLE 6

(*A*) $[M]_D^{20}$ *of homogeneous esters of* XCH_2COOH (*reference 36*)

X	DIPOLE MOMENT	k OF ACID DISSOCIATION	$[M]_D^{20}$ OF *l*-MENTHYL ESTER	$[M]_D^{20}$ OF *l*-OCTYL ESTER
$N(CH_3)_2$	+1.4	1.3×10^{-10} (?)	−156.9°	
H		1.8×10^{-5}	−157.3°	−11.8°
CH_3	+0.4	1.4	−160.2°	−13.0°
COOH	−0.9	160	−160.2°	
OC_2H_5		23	−160.6°	
OCH_3	−1.2	33	−165°	−16.3°
OH	−1.7	15	−165° (94°C.)	
Br	−1.5	138	−169°	−28.8°
Cl	−1.5	155	−171°	−17.9°
CN	−3.8	370	−174°	

(*B*) $[M]_D$ *of benzene solutions of* $RCH{=}NCH(C_6H_5)C_{10}H_6OH$ *derivatives* (*reference 3*)

ALDEHYDE USED TO FORM =CHR	$[M]_D$	$k \times 10^5$ (25°C.) OF CORRESPONDING ACIDS	LOG ($k \times 10^6$)
p-Dimethylaminobenzoic	+2676.0	0.94	0.974
p-Oxybenzoic	+1049.5	2.9	1.463
3-Bromo-*p*-oxybenzoic	+648.0		
Protocatechuic	+588.8	3.3	1.519
3-Nitroanisic	+559.6		
m-Toluic	+504.5	5.6	1.749
Benzoic	+373.1	6.6	1.820
m-Oxybenzoic	+362.6	8.33	1.920
p-Chlorobenzoic	+311.8	9.3	1.969
m-Bromobenzoic	+280.9	13.7	2.137
m-Chlorobenzoic	+255.9	15.5	2.190
m-Nitrobenzoic	+167.6	34.8	2.541
Salicylic	−85.7	106	3.025
o-Chlorobenzoic	−128.4	132	3.121
o-Bromobenzoic	−308.2	145	3.161
o-Nitrobenzoic	−990.7	657	3.826

alter the potential energy of a proton on the carboxyl group. They thus find that

$$\log k' \quad \frac{e\mu \cos \theta}{2.303kTDR^2} + \text{const.} \tag{48}$$

where k' is the dissociation constant of the acid under consideration, μ is the dipole moment of a substituent in the molecule, D is the effective dielectric constant of the medium between the dipole and the carboxyl (D depends on the shape of the molecule and the point of attachment of the dipole as well as on the nature of the molecule and its surrounding solvent), R is the distance from the dipole to the carboxyl, θ is the angle between the axis of the dipole and the direction of the carboxyl from the dipole, e is the charge of the proton, k is the Boltzmann constant, and T is the temperature.

Now the factor $\frac{\mu \cos \theta}{DR^2}$ in equation 48 gives the potential due to a dipole; in section III C we have seen that such a potential can influence the optical rotation caused by a given chromophore by an amount proportional to $\frac{\mu}{D} \cos \theta f(R)$, where D is assumed constant for points in the neighborhood of the chromophoric group, where R is large enough that $\cos \theta$ is the angular dependence function, and where $f(R)$ is the radial dependence function of the perturbation acting on the group. Therefore, we may write

$$\Omega = \Omega_0 + \frac{e\mu}{D} \cos \theta f(R) \tag{49}$$

For any given value of R (that is, for any series of ortho-, meta-, or para-derivatives, in the example under consideration) we may combine equations 48 and 49 to obtain

$$\Omega = \Omega_0 + 2.303kTR^2 f(R)(\log k' - \text{const.})$$

or

$$\Omega = \Omega_0' + 2.303kTR^2 f(R) \log k' \tag{50}$$

That is, there should be a linear relationship between the logarithm of the dissociation constant of all derivatives of an acid having substituents in the same position and the rotivities of derivatives of the acid. Furthermore, the plots of Ω *versus* $\log k'$ for all sets of derivatives must intersect at the point given by the derivative for which all $\mu = 0$ (in this case for the benzoic acid derivative).

In figure 15 are given plots of the rotation (which is here very nearly proportional to the rotivity, since dilute solutions of benzene were used for all measurements, so that the refractive index is nearly constant) against the logarithms of the dissociation constants of the acids related to the active molecules. The agreement is satisfactory throughout, considering the approximations made. It is interesting that the ortho-

derivatives give the poorest agreement, and it is just these for which we should expect the largest disturbance by the substituents acting as chromophores and also by their exerting vicinal actions different in nature from those pictured here, on account of their relative closeness to the asymmetric center of the molecule.

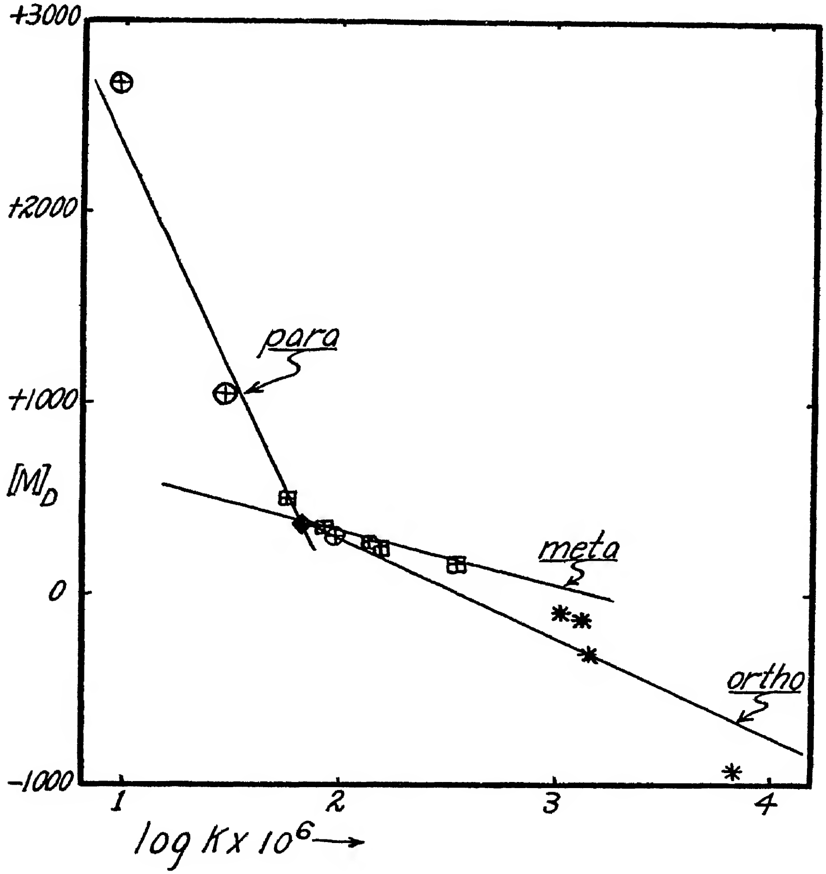


FIG. 15. Relation between optical rotation of $RCH{=}NCH(C_6H_5)C_{10}H_6OH$ and strengths of the acids RCOOH (data by Betti). ⊕, *p*-benzoic acid derivatives; ⊞, *m*-benzoic acid derivatives; *, *o*-benzoic acid derivatives; ◆, benzoic acid.

The data of Rule are also plotted (figures 16 and 17). Here the distances R are very nearly equal, while the angles θ probably are not too widely different, so that a plot both of dipole moment *versus* the rotation and of acid strength *versus* the rotation should give fairly straight lines. The other complicating factors mentioned above, however, make any

agreement which might exist obscure. In addition, the effects which are produced are not very large as compared with those obtained by Betti.

If the values for X = COOH and OH are neglected (these are marked with circles in the figures), the agreement becomes better. This may be caused by auxiliary disturbances from hydrogen-bonding between two active molecules, which are only possible in these compounds. Furthermore,

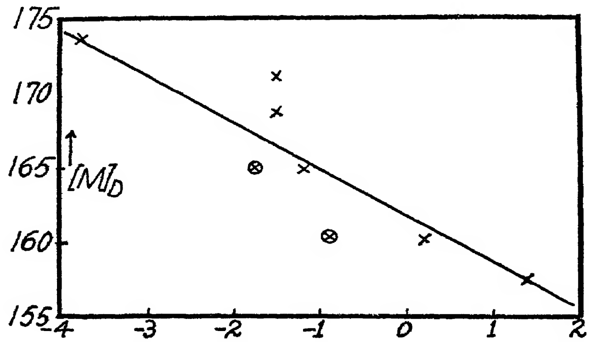


FIG. 16. Effect of bond moment of substituent on rotations of *l*-menthyl esters of XCH_2COOH

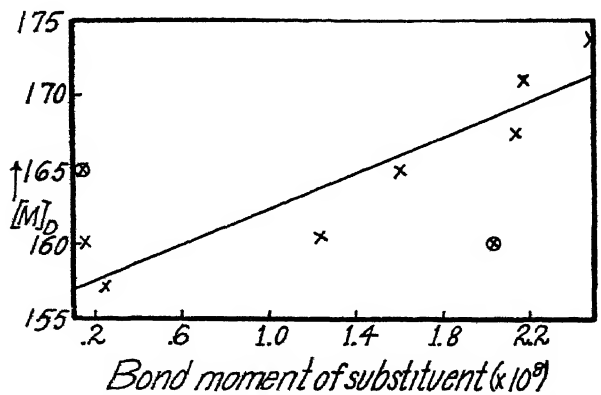


FIG. 17. Effect of acidity of XCH_2COOH on rotation of its *l*-menthyl ester

the rotation of the OH derivative was not taken at the same temperature as the others. On the whole, therefore, it may be said that these data tend to verify the ideas presented here.

E. OPTICAL SUPERPOSITION

van't Hoff's theory of the asymmetric carbon atom as a structural condition giving rise to the ability to rotate the plane of polarized light led him to propose an additivity relationship known as the principle of

optical superposition. According to this, in a molecule containing several asymmetric centers, each center contributes to the optical rotation independently of the others. Such a principle would be expected to be valid whenever the centers involved are widely enough separated; let us, however, examine a sugar, in which the centers are close to one another.

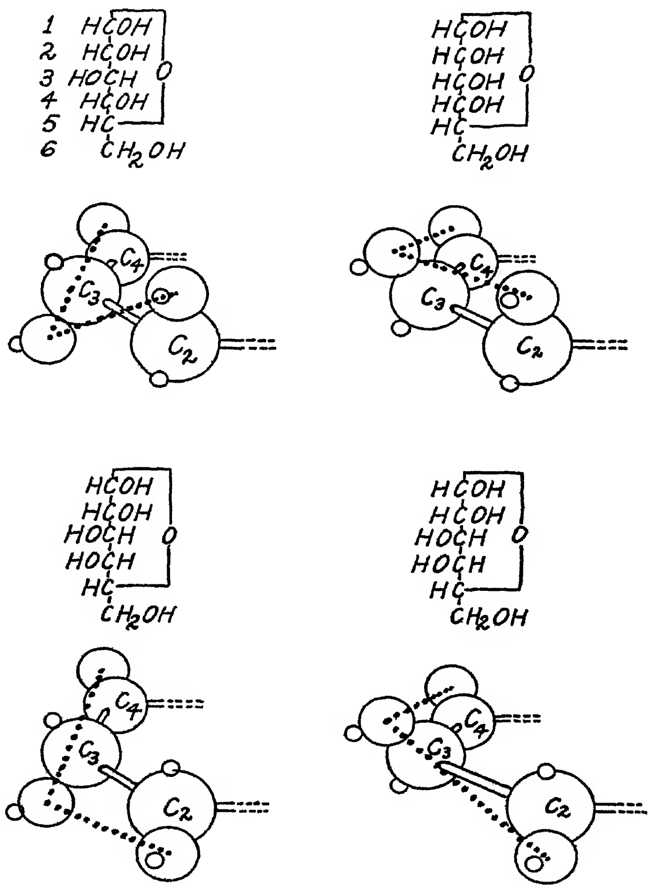


FIG. 18

It is readily seen from figure 18 that when the configuration about one center is changed, the changes occurring in the various inter-group distances and directions (these being indicated by the heavy dotted lines) depend very much on the configurations of the neighboring asymmetric carbon atoms. This is the more marked if the ring is puckered rather than planar.

For open-chain sugar derivatives a change of configuration about one center may not only bring about changes such as the above, but may also cause the chain to tend to assume a new conformation in order to adjust itself to the altered steric forces which are involved. Thus (see figure 19) if hydroxyl groups are assumed to repel each other very strongly, the chain will be found to be forced to assume different shapes for each diastereoisomer. This means that the rotation of a given isomer will depend on the configurations about all centers taken as a whole in contradiction to van't Hoff's principle of superposition. Therefore, it is not surprising that open-chain sugar derivatives fail to obey the principle of

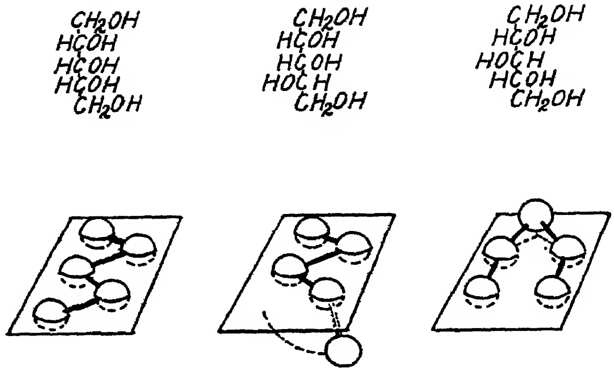
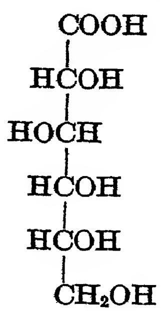
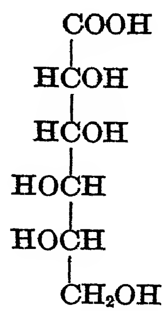


FIG. 19. Form of carbon chains if hydroxyls on neighboring atoms repel one another very strongly

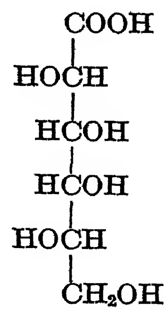
superposition. For instance, we consider the case of four sugar-acids cited by Freudenberg and Kuhn (11):



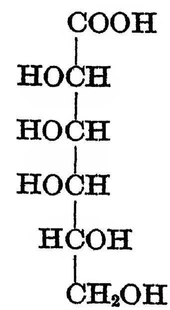
d-Gluconic acid
$[M]_D = -13°$



l-Mannonic acid
$[M]_D = +1°$



l-Galactonic acid
$[M]_D = +24°$



d-Talonic acid
$[M]_D = +33°$

The rotations of these acids, according to the principle of superposition, should be made up of the following contributions by each asymmetric atom:

+A	+A	−A	−A
−B	+B	+B	−B
+C	−C	+C	−C
+D	−D	−D	+D

Now it is apparent that the principle of superposition would predict the sum of the rotations of these compounds to be zero. It is actually +45°,—even larger than any of the individual rotations.

An instance in which the principle of superposition is probably obeyed because of the large distance between asymmetric centers is found in the dimenthylurethans of the diethyl tartrates (41):

$$\begin{array}{c} C_{10}H_{19}NHCOOCHCOOC_2H_5 \\ | \\ C_{10}H_{19}NHCOOCHCOOC_2H_5 \end{array}$$

l-Menthylurethan of diethyl *d*-tartrate: $[\alpha]_D = -74.34°$

l-Menthylurethan of diethyl *l*-tartrate: $[\alpha]_D = -38.76°$

$-57.78°$ (mean)

l-Menthylurethan of diethyl mesotartrate: $[\alpha]_D = -56.55°$

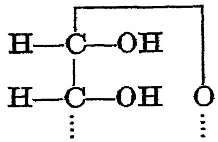
Similar results are found for the D-*l*-fenchyl urethans of the tartaric acids.

In these substances it would appear that the distance between the asymmetric centers in the menthyl radicals and those in the tartrate portion of the molecule is so great that the interaction is negligible. In taking the mean of the *d*- and *l*-esters, the effect of the tartrate portion cancels out, while in the meso compound the contribution is nil, so that the same residual rotations remain in both cases.

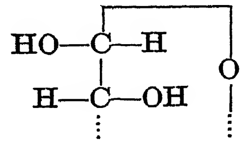
It might be surprising in the light of what has been said above that superposition appears to hold very well for many of the closed-ring sugars and their derivatives. This has led to the formulation of the so-called Hudson's rules of isorotation, whereby constant differences are found between the rotations of isomers which differ only in the configuration about a given asymmetric center. There are exceptions to these rules, particularly in the mannose, talose, and ketose series.

Freudenberg and Kuhn (11) explain both the occurrence of and the exceptions to Hudson's rules by assuming that groups attached to atoms twice removed from a given center are too far distant to be able to act appreciably in contributing to the rotation by the groups directly attached

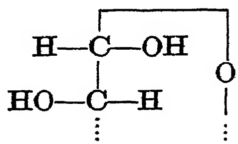
to the given center. Then the differences in the rotations of sugars containing the configurations



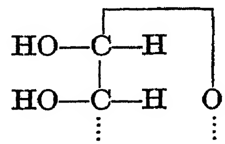
and



should be constant, regardless of the nature of the rest of the molecule. Furthermore, the value of this constant will not be the same as that obtained in the series



and



Since mannose and talose are of the second type, while the other sugars which are known to obey Hudson's rules belong to the first, this explanation is convincing.

The explanation becomes less convincing, however, when actual models of the sugars are made. If tetrahedral angles are maintained, it is seen that the groups on an atom twice removed from a given atom are in very close proximity to one or the other of the groups on the given atom. Besides, a large number of conformations of the ring are possible, and the stabilities of these would be expected to depend on the orientations of each and every group in the molecule, and so in turn affect the optical rotation to a marked degree. Furthermore, although vicinal actions fall off with distance, they probably do not fall off so rapidly as to have no more noticeable effect than they do on the validity of Hudson's rules. Therefore, it would appear that the problem deserves further consideration.

No final answer can be said to have been given, but some more general conditions under which Hudson's rules or superposition in general would hold have been developed (12). These are as follows: (1) that the vicinal actions of groups on a chromophore be additive (that is that the effect of one group on a chromophore is independent of the presence of other groups), and (2) that the structure to which the various groups whose configurations are changed are connected be symmetrical on the average (that is, that simultaneous change of configuration about two centers shall result merely in a change of sign of the vicinal actions between the groups involved in the change). If these conditions are satisfied, then regardless of how rapidly vicinal actions decrease with distance, there will appear to be superposition in the van't Hoff sense.

If it is these conditions which control the validity of Hudson's rules in the sugars, then some interesting conclusions can be drawn as to the structure of the rings involved. Thus, either the ring is flat, or else, if puckered in any way, it must exist in equal amounts in two conformations for each type of puckering which it can exhibit, so that the condition of the existence of the requisite effective symmetry in the ring is met. Presumably deviations from superposition occur whenever this is not true, and also when the additivity condition is violated (which is possible whenever there can be restricted rotation about the bonds of the attached groups, since a third group can then influence the vicinal actions between two other groups by influencing the position of maximum stability of one or the other of these groups about its axis).

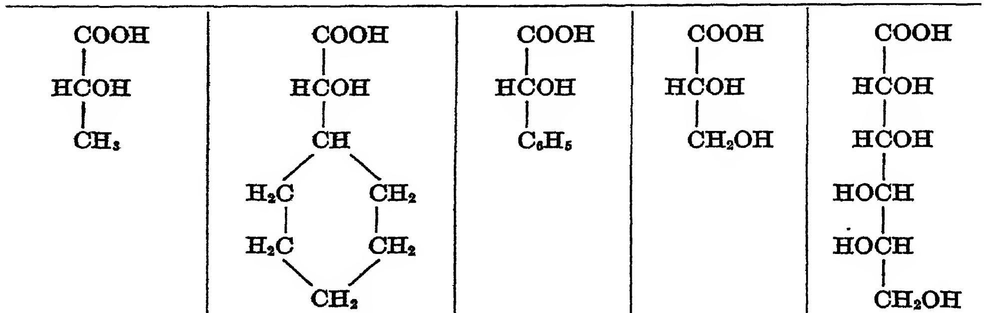
F. OPTICAL DISPLACEMENT

Consider the differences between the rotations of configurationally related α-hydroxy acids, RCHOHCOOH, and their amides, $RCHOHCONH_2$, where R may be any group (CH_3, C_2H_5, C_6H_5, etc.). The interactions which we must consider in calculating these differences are as follows:

(*a*) Interactions involving R and OH as chromophores and $CONH_2$ and COOH as vicinal groups. Since the COOH and $CONH_2$ groups are not very different, we would not expect them to act very differently as vicinal groups. (Thus, on the Kuhn theory and the Kirkwood theory, the similarity is expressible by similar orientations of the axes of the oscillators in the groups and similar values of the force constants and polarizabilities, while in the one-electron theory the charge distributions and dipole fields around the two groups are about the same, both groups having the same number of electrons and similar structures.) Therefore, if the substitution of $CONH_2$ for COOH produces no marked effect on the spatial structure of the group R (especially with respect to rotation about bonds, etc.), we should expect these interactions to remain about the same when we replace COOH by $CONH_2$, regardless of the nature of R.

(*b*) Interactions involving COOH and $CONH_2$ as chromophores and R and OH as vicinal groups. Since these groups are likely to show some difference when they act as chromophores, it is here that we would expect the major effect on the optical activity to show itself. Furthermore, the direction of the change in rotation when we pass from the acid to the amide should be the same for all R groups which are sufficiently similar. Just what is meant by "sufficiently similar" depends upon the theoretical approach that is taken. The Kuhn and Kirkwood theories require that the orientations of the axes of the oscillators in the different groups be similar. The one-electron theory requires that, insofar as dipoles are important, the dipoles possess about the same directions in space, and that.

insofar as incomplete screening of atoms is important, the atoms merely lie in the same direction in space with respect to the group being changed. In actual practice, such groups as methyl, phenyl, and cyclohexyl appear to meet well enough whichever of these criteria are actually operative, so that they may be called similar in the present sense of that word; this, along with a general experimental proof of what has been presented above, is indicated by the data (11) given below. All of the substances shown are, of course, configurationally related.



$[M]_D$ of: lactic	hexahydromandelic	mandelic	glyceric	*l*-mannonic
acid −3, amide +20 } +23	−42, +65 } +107	−233, −146 } +97	−2, +70 } +72	+1, +34 } +33

The ten other known cases in which one can convert an α-*d*-hydroxy acid to its amide all show a change of sign in the same direction as the above.

The reasoning which has been carried out for this more or less specific set of substances may be generalized to give the *displacement rule* (which was first[5] proposed by Freudenberg and later explained by Freudenberg and Kuhn in essentially the same terms which we have used here): Differences in rotation between analogous derivatives of analogously constructed compounds are approximately the same in magnitude and direction (21).

This rule is of great importance in explaining many of the qualitative relationships found in the data on optical rotation. Freudenberg proposed it as a basis for finding the relative configurations of similar substances which could not be related in any other way. The manner in which lactic acid and alanine were thus related to one another is illustrated in figure 20. The parallelism in the trend of the rotations of derivatives of (+) lactic acid and (+) alanine shows them to be configurationally similar. The method has also been employed in establishing other configurational relationships, but care must obviously be exercised in order to avoid its application to too widely different sets of compounds.

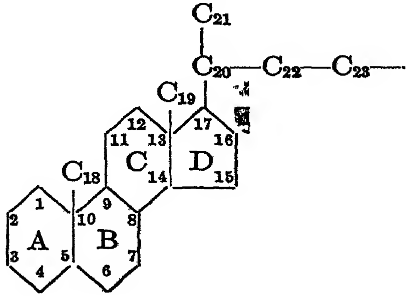
[5] Clough (7) in 1918 put forward a rule nearly identical with this.

In going from acid to neutral to basic solution, all amino acids probably undergo the following sequence of changes:

$$\begin{array}{ccccc} \begin{array}{c} \mathrm{H} \\ | \\ \mathrm{R{-}C{-}COOH} \\ | \\ \mathrm{NH_3^+} \end{array} & \longrightarrow & \begin{array}{c} \mathrm{H} \\ | \\ \mathrm{R{-}C{-}COO^-} \\ | \\ \mathrm{NH_3^+} \end{array} & \longrightarrow & \begin{array}{c} \mathrm{H} \\ | \\ \mathrm{R{-}C{-}COO^-} \\ | \\ \mathrm{NH_2} \end{array} \\ \text{(acid)} & & \text{(neutral)} & & \text{(basic)} \end{array}$$

Since the groups R are all fairly similar (in our sense of the word), we should expect to observe some regularity in the rotations which accompany these changes, provided all amino acids possess the same configuration. Lutz and Jergensons (28) have found such a regularity for natural amino acids, and their results are shown in figure 21. We may count this as the best sort of evidence that these naturally occurring amino acids possess the same configuration about their alpha carbon atoms. It is to be especially noted that in the details of the data shown there is considerable dependence on the nature of R, but that the general shapes of the curves are similar. This illustrates well the type of approximation which is embodied in the displacement rule.

Callow and Young (6) have pointed out some regularities in the rotations of the sterols which warrant attention here. The fundamental requirements which make these regularities possible seem to be (*1*) that vicinal actions are negligible beyond a certain distance, (*2*) that the steroid framework is spread out in space over a considerable area (so that any changes about position 17, for instance, have little effect on position 3), (*3*) that introduction of double bonds or a change in the *cis–trans* relationship of rings A and B does not result in any major movement of the atoms of the framework to new positions with respect to their more immediate neighbors, and (*4*) that when the effect of a given change at a certain point in the molecule is being examined, the changes which are allowed to occur at other points in the molecule are located at a sufficient distance from the given point as not to be able to exert any vicinal action on it.



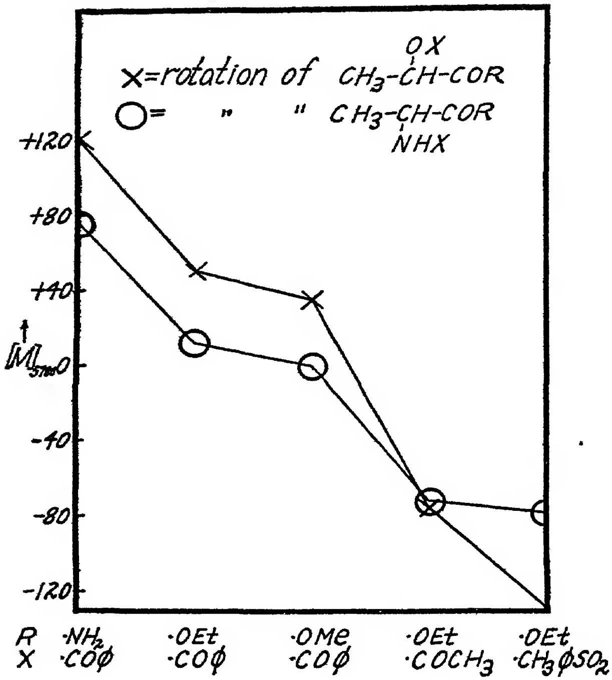


FIG. 20. Comparison of rotation of $CH_3CHOXCOR$ and $CH_3CH(NHX)COR$

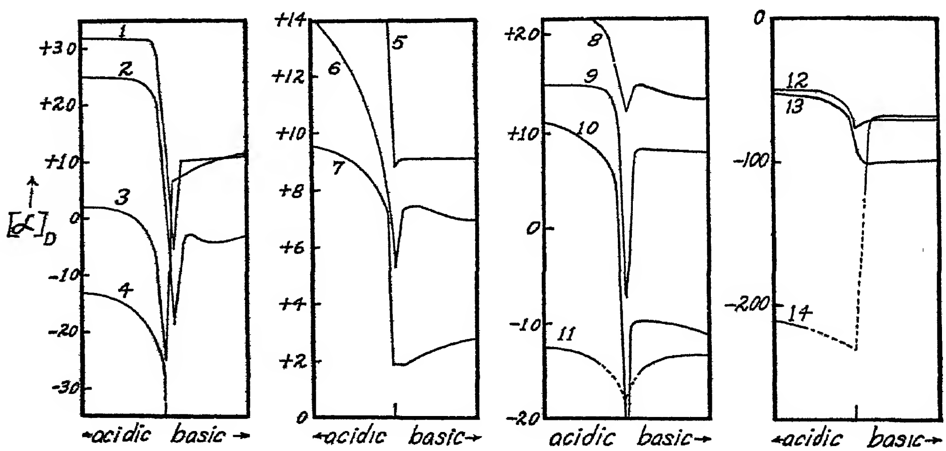


FIG. 21. Rotation of natural amino acids and acidity of their solutions. 1, glutamic acid; 2, aspartic acid; 3, tryptophan; 4, dihydroxyphenylalanine; 5, lysine; 6, ornithine; 7, alanine; 8, arginine; 9, leucine; 10, histidine; 11, tyrosine; 12, hydroxyproline; 13, proline; 14, cystine.

The observed regularities are of interest both in themselves and with regard to the exceptions which occur. They are therefore summarized below:

(*a*) A given change in configuration of a hydroxyl at position 3 for most epimers results in a change of optical rotation in the same direction. The magnitude of the change depends on the compound: for saturated sterols the change in molecular rotation is around 20°, while in the presence of a double bond the change[6] is larger, ranging from 50° to 400° depending on the proximity of the double bond to position 3.

(*b*) Introduction of a double bond at 4:5 increases the *d*-rotation by 150–200°. There are two exceptions, both involving the presence of a hydroxyl at 3: when this hydroxyl is *cis* to the methyl group, C_{18}, there is a decrease in *d*-rotation by 72° when the double bond is introduced, while when it is *trans* to C_{18} there is an increase in *d*-rotation by 373°. This discrepancy is readily explained: the 3-hydroxyl is very close to the position at which the double bond is introduced. Furthermore, in going from the *cis*-position to the *trans*, something close to reflection in one of the planes of symmetry of the double bond will be found to take place if a model is constructed. Therefore, the interaction of hydroxyl will be very large (say around 250°), and for the hydroxyl in the *cis*-position it depresses the normal change of +150° or 200° down to −50° or −100°, while for the *trans*-position it increases the normal change up to around +400°.

(*c*) Introduction of a double bond at 5:6 decreases the *d*-rotation by 200° or 300°. In one case in which there is already a double bond at 7:8, so that conjugation occurs when the new double bond is introduced, the decrease amounts to 436°. The bigger effect probably has largely to do with the movement of the absorption band nearer to the visible.

(*d*) Introduction of a double bond at 7:8 lowers the *d*-rotation by 62° in one case in which there is no conjugation, by 250–300° in two cases in which there is conjugation with a double bond at 5:6, and by 476° in one case in which there is conjugation with a double bond at 14:15. In two cases in which ring A is benzenoid (double bonds at 1:2, 3:4, 5:10) the introduction of a double bond at 7:8 results in an increase in *d*-rotation by 381° and 392°. Since the methyl at position 10 is no longer present, and since the benzene ring will act in an entirely different way as a chromophoric group, some such discrepancy is readily understandable.

(*e*) Introduction of a double bond at 8:14 is anomalous in its results on the optical rotation. This may be because in all cases the effect is small so that minor factors are able to throw the direction of change one way or the other.

[6] The fact that in the presence of double bonds at different positions the direction of change of rotation remains the same must be counted a matter of chance.

(*f*) Introduction of a double bond at 14:15 results in a small increase in *d*-rotation.

(*g*) Introduction of a double bond at 22:23 results in a decrease in *d*-rotation by 36° to 174°.

(*h*) Reduction of a carbonyl in position 17 to hydroxyl results in a decrease in *d*-rotation by about 200°.

(*i*) A few other operations on the sterol molecule appear to exert a regular effect on the rotation but too few data are available to draw safely any conclusions about them.

It should be clear from the discussion above that these regularities ought to be considered as illustrating the displacement rule rather than optical superposition, since they are then given a wider scope and the mediocre quantitative agreement which is found becomes less objectionable.

We take this opportunity to thank Dr. E. Gorin for many helpful discussions on the subject of optical rotatory power, and Dr. J. M. Sturtevant and Mr. K. L. Berry for sending us some of the results of their experimental researches bearing on section IV B in advance of publication.

REFERENCES

(1) Beckmann, C. O., and Cohen, K.: J. Chem. Phys. **4**, 784 (1936); **6**, 163 (1938).
(2) Berry, K. L., and Sturtevant, J. M.: J. Am. Chem. Soc. **61**, 3583 (1939).
(3) Betti, M.: Trans. Faraday Soc. **26**, 337 (1930).
(4) Bjerrum, N.: Z. physik. Chem. **106**, 219 (1923).
(5) Born, M.: Ann. Physik **55**, 177 (1918); Proc. Roy. Soc. (London) **A150**, 84 (1935).
(6) Callow, R. K., and Young, F. G.: Proc. Roy. Soc. (London) **A157**, 194 (1936).
(7) Clough, G. W.: J. Chem. Soc. **113**, 526 (1918).
(8) Condon, E. V.: Rev. Modern Phys. **9**, 432 (1937).
(9) Condon, E. V., Altar, W., and Eyring, H.: J. Chem. Phys. **5**, 753 (1937).
(10) Condon, E. V., and Shortley, G. H.: *The Theory of Atomic Spectra*. University Press, Cambridge (1935).
(11) Freudenberg, K., and Kuhn, W.: Ber. **64**, 703 (1931).
(12) Gorin, E., Kauzmann, W. J., and Walter, J. E.: J. Chem. Phys. **7**, 327 (1939).
(13) Gorin, E., Walter, J. E., and Eyring, H.: J. Chem. Phys. **6**, 824 (1938).
(14) Guye, P. A., and Aston, E. A.: Compt. rend. **124**, 194 (1897).
(15) Kirkwood, J. G.: J. Chem. Phys. **5**, 479 (1937).
(16) Kirkwood, J. G., Westheimer, F. H., and Shookhoff, M. W.: J. Chem. Phys. **6**, 506, 513 (1938); J. Am. Chem. Soc. **61**, 555, 1977 (1939).
(17) Kuhn, W.: Z. Physik **33**, 408 (1925).
(18) Kuhn, W.: Z. physik. Chem. **B4**, 14 (1929).
(19) Kuhn, W.: Trans. Faraday Soc. **26**, 293 (1930).
(20) Kuhn, W., and Gore, H. K.: Z. physik. Chem. **12**, 389 (1931).
(21) Kuhn, W., and Freudenberg, K.: In Eucken's *Hand- und Jahrbuch der chemischen Physik*, Vol. 8, part 3 (1932). Also Kuhn, W.: In Freudenberg's *Stereochemie*, p. 317, Hirschwaldsche Buchhandlung, Berlin (1932–33).

(22) KUHN, W., AND LEHMANN, H. L.: Z. Elektrochem. **37**, 549 (1931).
(23) LEVENE, P. A.: In Gilman's *Organic Chemistry*, Vol. II, p. 1779. John Wiley and Sons, Inc., New York (1938)..
(24) LIFSCHITZ, J.: Z. physik. Chem. **105**, 27 (1923); **114**, 485 (1925).
(25) LOWRY, T. M.: *Optical Rotatory Power*, Chap. IX. Longmans, Green and Company, London (1935).
(26) LOWRY, T. M., AND ALLSOPP, C. B.: Proc. Roy. Soc. (London) **146**, 313 (1934).
(27) LUCAS, R.: Trans. Faraday Soc. **26**, 418 (1930).
(28) LUTZ, O., AND JERGENSONS, B.: Ber. **63**, 448 (1930); **64**, 1221 (1931).
(29) MARGENAU, H.: Rev. Modern Phys. **11**, 1 (1939).
(30) MULLIKEN, R. S.: J. Chem. Phys. **3**, 564 (1935).
(31) OSEEN, C. W.: Ann. Physik **48**, 1 (1915).
(32) PATTERSON, T. S.: J. Chem. Soc. **109**, 1139 (1916).
(33) PENNY, W. G., AND SUTHERLAND, G. B. B. M.: J. Chem. Phys. **2**, 492 (1934).
(34) ROSENFELD, V. L.: Z. Physik **52**, 161 (1928). See also BORN, M., AND JORDAN, P.: *Elementare Quantenmechanik*, p. 250. J. Springer, Berlin (1930).
(35) RULE, H. G., *et al.*: A series of papers on solvent action appearing in the Journal of the Chemical Society from 1931 on.
(36) RULE, H. G.: Trans. Faraday Soc. **26**, 321 (1930).
(37) RULE, H. G., AND CHAMBERS, A. R.: J. Chem. Soc. **1937**, 145.
(38) SMYTH, C. P., ENGEL, E. W., AND WILSON, E. B., JR.: J. Am. Chem. Soc. **51**, 1736 (1929).
(39) SPONER, H.: *Molekülspektren*, Vol. I. J. Springer, Berlin (1935).
(40) THOMAS, L.: Naturwissenschaften **13**, 627 (1925).
(41) TSCHUGAEFF, L., AND GLEBKO, A.: Ber. **46**, 2752 (1913).
(42) VAN'T HOFF, J. H.: *Arrangement of Atoms in Space*, 2nd edition, p. 146. Vieweg und Sohn, Braunschweig (1898).
(43) WALDEN, P.: Ber. **38**, 363 (1905).
(44) WINTHER, C.: Z. physik. Chem. **41**, 161 (1902).
(45) WOLF, K. L., AND VOLKMANN, H.: Z. physik. Chem. **B3**, 139 (1929).

VALENCE RELATIONS AMONG THE METAL CARBONYLS[1,2]

ARTHUR A. BLANCHARD

Research Laboratory of Inorganic Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts

Received December 9, 1939

Nickel carbonyl has frequently been described as a pure coördination compound, or as a compound in which the nickel has a valence of zero. Such a statement is very vague. Any classification of the metal carbonyls in terms of valence must rest on a very precise definition of how the valence is computed, and if it so happens that the basis for the particular method of computing the valence is questionable, then the classification according to valence is also of doubtful value.

On the other hand, the use of the effective atomic number (E.A.N.) in classifying the volatile carbonyls, nitrosyl carbonyls, and carbonyl hydrides is extremely helpful. This concept is easily and clearly defined, and the properties of the carbonyls show a marked uniformity according to this classification.

The effective atomic number is the total number of electrons held within the sphere of the atom, and includes those that the atom itself furnishes, those that are added through electron transfer, and those that are added through the establishment of covalent and coördinate bonds. Whenever the effective atomic number of the central metal atom of the compound is equal to that of an inert gas, then it is possible for the compound to be volatile. By volatile we may somewhat arbitrarily say that we mean sufficiently volatile at temperatures below 100°C. to permit molecular weight determinations by the vapor density method.

Volatility connotes lack of cohesion between molecules, that is, lack of external field. The field of the molecules of the volatile carbonyls is very much self-contained just as that of the inert gases which have the same effective atomic number as the central atoms. It is, of course, obvious that the carbonyl molecule as a whole can lack external field only when

[1] Presented at the Symposium on Elements in Unusual States of Valence, held at the Ninety-eighth Meeting of the American Chemical Society, Boston, Massachusetts, September, 1939.

[2] Contribution No. 79 from the Research Laboratory of Inorganic Chemistry of the Massachusetts Institute of Technology.

the net electric charge is zero and when the field of the groups surrounding the central atom is also closed inside the molecule. Thus, all of the atoms of the molecule as well as the central metal atom must have the effective atomic number of an inert gas, and there must be a high degree of symmetry in the structure of the molecule.

KINDS OF VALENCE

Valence is one of the most fruitful concepts of the chemist. It is also the one which has caused the most uncertainty and controversy.

Its first service was in clarifying the distinction between atomic and equivalent weights. At that time no cognizance was taken of the polarity of valence, and no postulate was employed beyond that of the atomic theory. In fact, valence was hardly more than a number which described the experimental fact. Would that valence could continue to serve in this simple fashion!

The knowledge of the hydrocarbons, however, rendered the whole-number concept of valence untenable without the employment of new postulates of molecular structure.

The development of electrochemistry necessitated the assignment of polarity to valence, and thereupon arose the contradictions between polar and non-polar valence which seemed impossible of reconciliation until after the postulates of the electronic constitution of atoms had been introduced. G. N. Lewis offered the brilliant suggestion that pairs of electrons could at the same time fill positions in the electron shells of two atoms and thus raise the effective atomic number of each atom to that of an inert gas.

In forming compounds or ions of elements of the first two periods or of elements not more than two or three removed from an inert gas in the long periods of the periodic series, it is invariably the rule that each atom acquires the E.A.N. of an inert gas through either sharing or transfer of electrons. In the case of transfer, the valence is without question polar.

But the metals and metalloids in the middle of the long periods are too far removed from the inert gases to acquire easily the E.A.N. of the latter. Such elements are prone to acquire what is usually regarded as a polar valence of +2 or +3 through the loss of two or three electrons. This condition is stabilized through the acquisition of an outer shell of electrons around the thus incomplete kernel of the atom. This shell consists of lone pairs of electrons donated by coördinating molecules. Shells of eight and twelve held thus between the positive kernel of the central atom and the positive cores of the coördinating molecules are possessed of a remarkable stability. These shells correspond to coördination numbers of four and six, respectively; other coördination numbers are also recognized.

$$Ni \rightarrow Ni^{++} + 2e$$

$$Ni^{++} + 4NH_3 \rightarrow Ni\cdot 4NH_3^{++}$$

$$Ni^{++} + 4CN^- \rightarrow Ni(CN)_4^{--}$$

In these complexes the E.A.N. of nickel is 34.

$$Co \rightarrow Co^{+++} + 3e$$

$$Co^{+++} + 6NH_3 \rightarrow Co\cdot 6NH_3^{+++}$$

$$Co^{+++} + 6CN^- \rightarrow Co(CN)_6^{---}$$

In these complexes the E.A.N. of cobalt is 36, the same as the atomic number of krypton.

The stability of the coördinating sphere of electrons is the determining factor in the formation of the complex ions, and thus in many cases the E.A.N. may be different from that of an inert gas, and the inner electron layers of the central atom may be incomplete. However, in a surprisingly large number of cases the E.A.N. of the central atom is actually the same as that of an inert gas, and it is a fair assumption that the attainment of this E.A.N. is a factor which increases the stability.

Carbon monoxide, nitric oxide, and the cyanide ion have very similar complex-forming functions (63). Uncombined their electronic structures may be represented as

:C:::O: [:C:::N:]⁻   and   :N:⋮:O:

The extra electron which nitric oxide possesses over carbon monoxide is so completely imprisoned in the structure that the physical properties of the two substances—as well as of nitrogen :N:::N:—are very similar.[3] Nevertheless, this extra electron does have some tendency to escape, leaving the positive radical [:N:::O:]⁺. The existence of solid crystalline nitrosyl perchlorate, $NOClO_4$ (12), nitrosyl fluoborate, $NOBF_4$ (64), and nitrosylsulfuric acid, $NO(OSO_2OH)$, and particularly the fact that $NOClO_4$ dissolved in methyl nitrate has the characteristic conductivity of a salt (8) all support this view of the positive character of the coördinating nitrosyl radical.

In linking up in complexes there are two possibilities of electron arrangement:

(I) M:C:::O: M:N:::O:

(II) M::C::O::   M::N::O::

[3] The complete absence of tendency of the nitrogen molecule to take part in complex building must be ascribed to the complete symmetry of the positive charges.

In case I the carbon monoxide has established one coördinate bond with the metal atom and has thus added two electrons to increase the E.A.N. The nitric oxide has established one coördinate bond with the metal atom, and the imprisoned electron has been completely transferred to the metal atom; thus three electrons have been added to increase the E.A.N.

In case II, two covalent bonds have been established between the carbon monoxide and the metal atom, but the metal atom has furnished two of the electrons so that the increase in the E.A.N. is the same as in case I,—namely, two. The nitric oxide also has increased the E.A.N. of the metal atom by three, the same as in case I.

It is interesting to note that electron diffraction studies (3, 4) of nickel carbonyl, cobalt nitrosyl carbonyl, $Co(NO)(CO)_3$, and iron nitrosyl carbonyl, $Fe(NO)_2(CO)_2$, have indicated tetrahedral structure, a straight-line arrangement of M—C—O and M—N—O, and bond distances which are in agreement with the supposition of a condition of resonance between arrangements I and II.

Likewise, the carbonyl hydrides $Co(CO)_4H$ and $Fe(CO)_4H_2$ have been subjected to electron diffraction studies by Ewens and Lister (7), with the conclusion that the same condition of resonance exists between arrangements I and II and that the hydrogen atom is situated at the end of the chain M—C—O—H

$$\underset{\text{I}}{\mathrm{M:C:::O:H}} \qquad \underset{\text{II}}{\mathrm{M::C::\ddot{O}:H}}$$

Other electron diffraction studies indicate that the structure in iron pentacarbonyl is that of a trigonal bipyramid, and that in the hexacarbonyls the structure is octahedral (5).

In table 1 are listed the known monomeric carbonyls with their effective atomic numbers figured on the supposition that each carbon monoxide adds two, each nitric oxide adds three, and each hydrogen adds one (2). All of the carbonyls in this list are volatile according to the definition we have adopted.

The polymeric carbonyls, $Fe_2(CO)_9$, $[Fe(CO)_4]_3$, $[Co(CO)_4]_2$, $[Co(CO)_3]_4$, $Ru_2(CO)_9$, and $[Ru(CO)_4]_n$ are not volatile according to our definition, although $Fe_2(CO)_9$, $[Co(CO)_4]_2$, and $Ru_2(CO)_9$ can be sublimed.

In all of these cases, however, it is possible to assign to the metal atoms the E.A.N. of an inert gas if one is allowed to save the necessary number of electrons by the linking of carbonyl groups together, thus, MCO:OCM, or to bridge two metal atoms with a single carbonyl group M:CO:M and to choose either arrangement I or arrangement II when only one of them will fit.

It is to be noted that the volatile nickel carbonyl already satisfies the tendency for the central atom to acquire the E.A.N. of 36, and that this is the only carbonyl of nickel. On the other hand, cobalt and iron form polymeric carbonyls, nitrosyl carbonyls, and carbonyl hydrides in order to satisfy this tendency.

POLAR VALENCE

Polar valence is a convenient term to use, and there can be no misunderstanding of its significance when it is applied to the charge of a monatomic ion such as Na^+ or Cl^-. When, however, it is applied to the single atoms in complexes,—either neutral molecules or ions,—we frequently run into a good deal of uncertainty. The polar valence, or electrovalence, or oxidation and reduction number, is supposedly the number of faradays necessary to convert one atomic weight of the element at the electrode to the condition in question. Other constituents which enter the complex ion are supposedly unchanged electrochemically. Too much supposition is involved to make it possible to regard polar valence as an unqualified statement of experimental fact.

TABLE 1

Effective atomic numbers in the monomeric carbonyls[4]

CARBONYL	E.A.N.	INERT GAS
BH_3CO	5 + 3 + 2 = 10	Ne
$Cr(CO)_6$	24 + 12 = 36	Kr
$Fe(CO)_5$	26 + 10 = 36	Kr
$Fe(CO)_2(NO)_2$	26 + 4 + 6 = 36	Kr
$Fe(CO)_4H_2$	26 + 8 + 2 = 36	Kr
$Co(CO)_3NO$	27 + 6 + 3 = 36	Kr
$Co(CO)_4H$	27 + 8 + 1 = 36	Kr
$Ni(CO)_4$	28 + 8 = 36	Kr
$Mo(CO)_6$	42 + 12 = 54	Xe
$Ru(CO)_5$	44 + 10 = 54	Xe
$W(CO)_6$	74 + 12 = 86	Rn

[4] Since this paper was submitted, Hieber and Schulten (Z. anorg. allgem. Chem. **243**, 164 (December, 1939)) have announced the preparation of carbonyl halides of rhenium, $Re(CO)_5X$, where X = Cl, Br, or I. These compounds have a very low vapor pressure but can be readily sublimed in a current of carbon monoxide, the iodine compound at 90°C. and the chlorine compound at 140°C. In these compounds the rhenium has the E.A.N. of 86 (radon).

Chlorine and carbon monoxide both combine spontaneously with nickel to produce a new substance. Certainly we must go beyond these experimental facts to find a reason to give nickel a valence of +2 in the chloride

and 0 in the carbonyl. Chlorine will displace iodine from iodides because chlorine is said to be more electronegative than iodine, that is, the free chlorine takes the negative electric charge away from the iodine of the iodide.

But carbon monoxide is capable of displacing halogens from a number of halides; for example, when it is passed over ruthenium triodide at 250°C., free iodine is liberated.

$$RuI_3 + 2CO \rightarrow Ru(CO)_2I_2 + \tfrac{1}{2}I_2$$

Carbon monoxide is capable even of displacing bromine and chlorine in similar reactions. All of this shows that polar valence cannot always be stated in terms of experimental fact alone.

POLAR VALENCE IN THE VOLATILE CARBONYLS

Nickel carbonyl

Neutral carbon monoxide plus neutral nickel yields neutral nickel carbonyl. This has been described as a pure coördination compound. If the carbon monoxide molecule still holds all of its original electrons, it seems justifiable to call the nickel zero-valent according to the same reasoning that, e.g., copper is divalent in the $[Cu \cdot 4NH_3]^{++}$ complex.

Cobalt nitrosyl carbonyl

Cobalt nitrosyl carbonyl has an electronic structure identical with that of nickel carbonyl, but an electron transfer has occurred; the imprisoned electron of the nitric oxide has been completely transferred to the cobalt atom to complete the pattern of its kernel electrons. In counting the polar valence of cobalt it would seem as if this transferred electron should be counted, whereas the eight electrons of the coördinate bonds should be ignored. Thus the polar valence of cobalt would be −1.

Cobalt carbonyl hydride, $(CO)_3CoCO:H$

The electron of the neutral hydrogen atom has passed along the chain to take its place in the cobalt atom and again the polar valence of cobalt would be −1.

Dimeric cobalt tetracarbonyl

If neutral carbon monoxide is coördinatively added to cobalt, it might seem that the valence should be zero. However, there is undoubtedly the same shift of an electron into the kernel structure of the cobalt atom, the electrons being derived from the release of one pair when a bond such as $(CO)_3CoCO:COCo(CO)_3$ is established. This seems to be an argument that the valence is −1 even here.

Iron pentacarbonyl

If all five carbonyl groups are separately attached direct to the central iron atom, the latter should be assigned a valence of zero in accordance with the argument used with nickel carbonyl. If, however, the structure of iron pentacarbonyl is like that of nickel carbonyl with four carbonyl groups attached to the iron atom, and the fifth carbonyl group completing a ring

$$\begin{array}{cccc} \mathrm{CO:} & & \mathrm{:CO:} & \mathrm{C} \\ & \mathrm{Fe} & & \\ \mathrm{CO:} & & \mathrm{:CO:} & \mathrm{O} \end{array}$$

then the electrons released in creating the carbonyl–carbonyl bonds would be transferred completely to the iron instead of coördinatively donated, and the valence of the iron would be -2. Structure II alone, however, would fit this interpretation.

Iron nitrosyl carbonyl and iron carbonyl hydride

In $Fe(NO)_2(CO)_2$ and in $Fe(CO)_4H_2$ the valence of iron would be -2.

Trimeric iron tetracarbonyl

The same argument as with cobalt tetracarbonyl leaves the uncertainty whether to assign the valence 0 or -2 to the metal atom.

Hexacarbonyls of chromium, molybdenum, and tungsten

Here the probable octahedral arrangement with coördination number of six would indicate a valence of zero for the metal, whereas with two rings of three carbonyl groups each, i.e., with coördination number of four, the valence of -4 would be indicated.

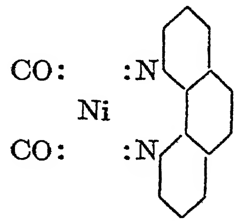
The foregoing arguments are confusing. The volatile carbonyls are so very non-polar in their character that a polar valence of the metal atom has little significance in any case. It seems to be the part of wisdom to avoid the use of polar valence with these compounds and to base our systematization of them on the concept of E.A.N. Only in the case of the hydrides, which behave as weak acids, we have the ions $Co(CO)_4^-$ and $Fe(CO)_4^{--}$, and to the ions as a whole of course a polar valence applies.

AMINE-SUBSTITUTED CARBONYLS

In the bewildering array of amine-substituted carbonyls (2), it is true in a great many of the cases that the amine molecules merely occupy the same number of coördinate positions as the displaced carbonyl groups, and thus the E.A.N. of the central atom is unchanged. For example,

$$Ni(CO)_4 + o\text{-phthr} \rightarrow Ni(CO)_2 \cdot o\text{-phthr} + 2CO$$

Here we have a chelate complex and the ortho-phenanthroline occupies two of the coördinate positions around the nickel atom.



It is noteworthy that no instance has been observed in which the carbonyl groups have been completely displaced by other neutral groups to form a pure coördination compound of the metal.

There is no record that any of the amine-substituted carbonyls displays any volatility, and no great surprise need be felt if the metal atom does not seem to possess the E.A.N. of an inert gas in a compound such as, for example, $Ni_2(CO)_3py_2$ (py = pyridine).

COMPLEX IONS CONTAINING CARBONYL AND NITROSYL GROUPS

The major influence upon the stability of a complex ion is undoubtedly the formation of a coördinating layer of electrons between the central atom and the surrounding addenda. The strength of this coördinating layer is enough to allow a considerable variation in the polar valence of the central atom; indeed some very unusual valences of the central atom have been noted in complex ions. Carbonyl and nitrosyl groups occur as addenda in a good many complex ions, and the method of coördination by means of a lone pair of electrons is entirely of the same nature as with other complex-forming groups, such as cyanide ion, chloride ion, and ammonia.

$$:C:::N: \qquad :Cl \qquad :NH_3$$

The volatile carbonyls are, in terms of the Werner coördination theory, just a special case in which the complex has a net charge of zero. Since an effective atomic number equal to that of an inert gas is an essential condition for the formation of a volatile carbonyl, it is of interest to study the relation of the E.A.N. to the stability of complex ions. The compilation in table 2 contains at least fairly well authenticated complex ions that contain carbonyl and nitrosyl groups, and for the sake of showing the trends, some other ions are included. In general, when the existence of a salt-like substance is reported which has, e.g., the composition $M_5Mn(CN)_6$, it is assumed that the complex $[Mn(CN)_6]^{-----}$ is indicated. (M represents a univalent alkali metal.)

The polar valence is figured as the difference between the atomic number

and the number of electrons inside (but not including) the coördinating shell. Carbon monoxide, :C:::O:, cyanide ion, :C:::N:, chloride ion

TABLE 2

Effective atomic numbers of the central atoms in complex ions

COMPLEX ION	POLAR VALENCE	E.A.N.	REFERENCE
$M_3[Cr(CN)_6]$	+3	24 − 3 + 12 = 33	Pascal (47)
$M_4[Cr(CN)_6]$	+2	24 − 2 + 12 = 34	Pascal (48)
$M_3[Mn(CN)_6]$	+3	25 − 3 + 12 = 34	Pascal (39)
$M_4[Mn(CN)_6]$	+2	25 − 2 + 12 = 35	Pascal (39)
$M_5[Mn(CN)_6]$	+1	25 − 1 + 12 = 36	Manchot and Gall (27)
$M_3[Mn(CN)_5NO]$	+1	25 − 1 + 12 = 36	Manchot and Schmid (35)
$M_3[Fe(CN)_6]$	+3	26 − 3 + 12 = 35	Pascal (42)
$M_4[Fe(CN)_6]$	+2	26 − 2 + 12 = 36	Pascal (41, 13)
$M_3[Fe(CN)_5CO]$	+2	26 − 2 + 12 = 36	Manchot and Woringer (38)
$M_2[Fe(CN)_5NO]$	+2	26 − 2 + 12 = 36	Manchot and Woringer (38)
$M_3[Fe(CN)_5NO]$	+1	26 − 1 + 12 = 37	Manchot and Woringer (38)
$M_3[Fe(CN)_5NH_3]$	+2	26 − 2 + 12 = 36	Hoffman (11)
$M_2[Fe(CO)_4]$	−2	26 + 2 + 8 = 36	Blanchard (2)
$M_3[Co(CN)_6]$	+3	27 − 3 + 12 = 36	Pascal (46)
$[Co(NH_3)_6]Cl_3$	+3	27 − 3 + 12 = 36	Pascal (45)
$M_4[Co(CN)_6]$	+2	27 − 2 + 12 = 37	Pascal (44)
$M[Co(CO)_4]$	−1	27 − 1 + 8 = 36	Blanchard (2)
$M_3[Co(CN)_5CO]$	+2	27 − 2 + 12 = 37	Manchot and Gall (26)
$M_2[Ni(CN)_4]$	+2	28 − 2 + 8 = 34	Pascal (43)
$M_2[Ni(CN)_3]$	+1	28 − 1 + 6 = 33	Burgess (6, 1)
$M_4[Ni(CN)_4]$	0	28 + 8 = 36	Burgess (6)
$M_2[Ni(CN)_3NO]$	0	28 + 8 = 36	Manchot (18)
$M_4[Mo(CN)_8]$	+4	42 − 4 + 16 = 54	Pascal (49)
$M_4[Ru(CN)_6]$	+2	44 − 2 + 12 = 54	Howe (13)
$M_2[Ru(CN)_5NO]$	+2	44 − 2 + 12 = 54	Manchot and Dusing (19)
$M_2[RuCl_5NO]$	+2	44 − 2 + 12 = 54	Manchot and Schmid (36)
$M_3[Rh(CN)_6]$	+3	45 − 3 + 12 = 54	Pascal (56)
$M_2[PdCl_6]$	+4	46 − 4 + 12 = 54	Pascal (60)
$M_2[Pd(CN)_4]$	+2	46 − 2 + 8 = 52	Pascal (61)
$M_4[W(CN)_8]$	+4	74 − 4 + 16 = 86	Pascal (50)
$M_2[OsCl_6]$	+4	76 − 4 + 12 = 84	Pascal (51)
$M_3[OsCl_6]$	+3	76 − 3 + 12 = 85	Pascal (53)
$M_2[OsCl_5NO]$	+2	76 − 2 + 12 = 86	Pascal (52)
$M_4[Os(CN)_6]$	+2	76 − 2 + 12 = 86	Pascal (54)
$M_3[Ir(CN)_6]$	+3	77 − 3 + 12 = 86	Pascal (55)
$M_2[PtCl_6]$	+4	78 − 4 + 12 = 86	Pascal (57)
$M_2[PtCl_2(CN)_4]$	+4	78 − 4 + 12 = 86	Pascal (59)
$M_2[Pt(CN)_4]$	+2	78 − 2 + 8 = 84	Pascal (58)

:Cl, and ammonia, $:NH_3$, each donates two electrons to the coördinate shell and does not affect the polar valence. Nitric oxide is assumed to

lose one electron entirely to the metal and thus lower the polar valence of the latter, the remaining positive nitrosyl group :N:::O: contributing two electrons to the coördinate shell.

It is noticed at once that in a great many of these complex ions the E.A.N. is actually that of an inert gas. In cases in which the E.A.N. is different there is almost invariably a manifest trend towards the stable number.

For example, in the complex iron cyanides the ferricyanide ion, $Fe(CN)_6^{---}$ (E.A.N. = 35), tends to gain an electron to change to the ferrocyanide ion (E.A.N. = 36). The replacement of one cyanide ion by a carbonyl group in the ferricyanide ion, thus forming the ion $Fe(CN)_5CO^{---}$, again raises the E.A.N. to 36. Introduction of ammonia to replace a cyanide ion likewise raises the E.A.N. to 36. The well-known nitroprusside ion, $Fe(CN)_5NO^{--}$, bears two negative charges, thus giving to the central atom the E.A.N. of 36. The triply charged ion, $Fe(CN)_5NO^{---}$, in which the E.A.N. is 37, can be prepared, but it is unstable and goes over spontaneously to the common nitroprusside (38).

Among the cobalt cyanides the cobaltocyanide ion, $Co(CN)_6^{----}$ (E.A.N. = 37), tends strongly to go over into the cobaltic cyanide ion, $Co(CN)_6^{---}$, with an E.A.N. of 36.

In preparing the carbonyl-substituted complex $Co(CN)_5CO^{---}$ (E.A.N. = 37) Manchot and Gall took extreme precautions to suppress the spontaneous evolution of hydrogen, which would have increased the valence of cobalt from 2 to 3. In all probability the ion $Co(CN)_5CO^{--}$ (E.A.N. = 36) would be formed if it were given a fair chance.

Among the nickel cyanides, the E.A.N. of nickel in the rather unstable $Ni(CN)_4^{--}$ is 34, but the tendency towards an E.A.N. of 36 is shown in the very unusual ion $Ni(CN)_4^{----}$, in which the polar valence of nickel is zero (6).

W. M. Burgess has found that when a solution of potassium nickelocyanide in liquid ammonia is treated with excess of potassium, a yellow precipitate of the composition $K_4Ni(CN)_4$ is formed; with the $K_2Ni(CN)_4$ in excess, however, a red precipitate of $K_2Ni(CN)_3$, previously described by Bellucci (1), is obtained. Burgess also has found that ferricyanides (E.A.N. = 35) are reduced to ferrocyanides (E.A.N. = 36), but that ferrocyanides are not reduced further by liquid ammonia solutions of sodium, potassium, or calcium.

Among the cyanides of manganese the trend is shown in the formation of $Mn(CN)_6^{-----}$ (E.A.N. = 36), a most unusual ion with five negative charges, and in the ion $Mn(CN)_5NO^{---}$, in which manganese also has the unusual polar valence of +1.

Among the cyanides of chromium only the chromo- and the chromi-

cyanide ions have been reported, but the trend would seem to point to the hexavalent ion $[Cr(CN)_6]^{------}$, in which chromium would be zerovalent.

Molybdenum and tungsten show, respectively, the E.A.N. = 54 of xenon and the E.A.N. = 86 of radon in the complexes $[Mn(CN)_8]^{----}$ and $[W(CN)_8]^{----}$, which have the unusual coördination number of eight.

Among the platinum metals the prevalence of the E.A.N. of xenon and radon is very obvious. With divalent palladium and platinum, as also with nickel, such an E.A.N. is structurally impossible, and we should look to see if compounds corresponding to $K_4[Ni(CN)_4]$ and $K_2[Ni(CN)_3NO]$ can be prepared.

TABLE 3

Carbonyl halides of heavy metals

COMPOUND	POLAR VALENCE	E.A.N.	REFERENCE
$CuCO\cdot X$	+1	29 − 1 + 4 = 32	(23)
$CuNO\cdot X_2$	+1	29 − 1 + 6 = 34	(14)
$Ru(CO)_2\cdot X_2$	+2	44 − 2 + 8 = 50	(28)
$RuCO\cdot Br$	+1	44 − 1 + 4 = 47	(21)
$PdCl_2\cdot CO$	+2	46 − 2 + 6 = 50	(31)
$PdCl_2\cdot 2NO$	0	46 + 8 = 54	(37)
$OsX_2\cdot 3CO$	+2	76 − 2 + 10 = 84	(29)
$IrCl_2\cdot 2CO$	+2	77 − 2 + 8 = 83	(24)
$PtCl_2\cdot CO$	+2	78 − 2 + 6 = 82	(17, 62)
$PtCl_2\cdot 2CO$	+2	78 − 2 + 8 = 84	(17, 62)
$AuCl\cdot CO$	+1	79 − 1 + 4 = 82	(25)
$Fe(NO)_2I$	−1	26 + 1 + 6 = 33	(9, 22)
$Co(NO)_2I$	−1	27 + 1 + 6 = 34	(10)

CARBONYL HALIDES

Carbon monoxide and nitric oxide combine with the salts, notably the halides, of the heavy metals, forming compounds which vary a good deal in character but in general are sublimable (although not volatile in terms of our arbitrary standard), lack pronounced salt-like character, are soluble in organic solvents, are hydrolyzed or decomposed by water, and possess unusual valences of the metal atom.

The valence and the E.A.N. in table 3 are figured on the basis of the monomeric formula and on the assumption that nitric oxide has transferred one electron and donated two electrons to the metal atom. The lack of volatility and for the most part the absence of molecular weight determinations of any kind deprives the E.A.N. numbers of significance,

because it is impossible to estimate how many electrons are gained by CO or NO bridges between metal atoms. Further compounds, in which no attempt is made to assign a valence or an E.A.N., add still more to the confusion of the picture: $Ag_2SO_4 \cdot CO$ (16, 32), $RhCl_2 \cdot RhO \cdot 3CO$ (30), $RhCl_2 \cdot RhO \cdot 3NO$ (34), $2PtCl_2 \cdot 3CO$ (17, 62), $FeSO_4 \cdot NO$ (40), $Fe(NO)_4$ (20), $Ru(NO)_{4-5}$ (33), and $COHg(OC_2H_5)(COOCH_3)$ (15).

The only sure generalization which can be made is that carbonyl and nitrosyl groups have a strong tendency to coördinate with heavy-metal atoms and that in general such coördination enhances the stability of an unusual valence state of the metal.

It probably is unwarranted to assume that always one electron is completely transferred to the metal when nitric oxide coördinates, and hence the polar valences derived on that basis should be taken with some reservation.

It is interesting to note that in the similar pair of nitrosyl compounds $Fe(NO)_2I$ and $Co(NO)_2I$ the atomic number of the metal has no effect in determining the structure of the complex, whereas in the pair $K_3[Mn(CN)_5NO]$ and $K_2[Fe(CN)_5NO]$ the tendency to acquire the E.A.N. of krypton is sufficient to determine the polar valence of the metal and the charge of the complex ion.

In the pair of compounds $RhCl_2 \cdot RhO \cdot 3CO$ and $RhCl_2 \cdot RhO \cdot 3NO$ the carbonyl and nitrosyl groups are interchangeable (34), and either nitric oxide or carbon monoxide is able to replace the other without breaking the structure of the compound.

SUMMARY

The use of the term "valence" involves many uncertainties unless the method of computing the valence is very carefully detailed in every case.

This is particularly true when we are dealing with the volatile carbonyls. Indeed any method of computing the valence with these compounds is so involved in uncertain hypotheses that it would seem better to avoid altogether the use of the term.

The concept of effective atomic number (E.A.N.), however, proves an excellent basis for classifying the volatile carbonyls.

In the formation of complex ions containing carbonyl and nitrosyl groups, two factors are apparent: first, the tendency to form a stable coördinating layer of electrons between the central atom and the surrounding addenda; second, the tendency of the central atom to acquire the E.A.N. of an inert gas. It is unmistakable that the ions in which this second tendency has also functioned are more stable than those in which the first tendency alone has been satisfied.

The non-volatile or little-volatile carbonyl and nitrosyl halides and salts

of the heavy metals present a very incomprehensible picture. Too little is known of their properties to draw any valid conclusions. In general, the coördination of carbonyl and nitrosyl groups appears to stabilize the otherwise less stable polar valence of the heavy metal. But even this generalization is subject to the uncertainty of the definition of the polar valence.

REFERENCES[5]

(1) Bellucci, I. and Corelli, R.: Z. anorg. Chem. **86**, 88 (1914).
(2) Blanchard, A. A.: Chem. Rev. **21**, 3 (1937).
(3) Brockway, L. O., and Anderson, J. S.: Trans. Faraday Soc. **33**, 1233 (1937).
(4) Brockway, L. O., and Cross, P. C.: J. Chem. Phys. **3**, 828 (1935).
(5) Brockway, L. O., Ewens, R. V. G., and Lister, M. W.: Trans. Faraday Soc. **34**, 1350 (1938).
(6) Burgess, W. M.: Paper read at the Symposium on Elements in Unusual States of Valence, held at the Ninety-eighth Meeting of the American Chemical Society, Boston, September, 1939.
(7) Ewens, R. V. G., and Lister, M. W.: Trans. Faraday Soc. **35**, 681 (1939).
(8) Hantzsch, A., and Berger, K.: Z. anorg. allgem. Chem. **190**, 321 (1930).
(9) Hieber, W., and Anderson, J. S.: Z. anorg. allgem. Chem. **211**, 132 (1933).
(10) Hieber, W., and Marin, R.: Z. anorg. allgem. Chem. **240**, 241 (1939).
(11) Hoffman, K. A.: Z. anorg. Chem. **10**, 264 (1895); Ann. **312**, 1 (1900).
(12) Hoffman, K. A., and Zedwitz, G. A.: Ber. **42**, 2031 (1909).
(13) Howe, J. L.: J. Am. Chem. Soc. **18**, 981 (1896).
(14) Manchot, W.: Ber. **47**, 1601 (1914).
(15) Manchot, W.: Ber. **53**, 984 (1920); **54**, 571 (1921).
(16) Manchot, W.: Ber. **57**, 1157 (1924).
(17) Manchot, W.: Ber. **58**, 2518 (1925).
(18) Manchot, W.: Ber. **59**, 2448 (1926).
(19) Manchot, W., and Dusing, J.: Ber. **63**, 1226 (1930).
(20) Manchot, W., and Enk, E.: Ann. **470**, 275 (1929).
(21) Manchot, W., and Enk, E.: Ber. **63**, 1635 (1930).
(22) Manchot, W., and Fischer, H.: Dissertation, Technische Hochschule, München, 1937; Z. anorg. allgem. Chem. **240**, 242 (1939).
(23) Manchot, W., and Friend, J. N.: Ann. **359**, 100 (1908).
(24) Manchot, W., and Gall, H.: Ber. **58**, 232 (1925).
(25) Manchot, W., and Gall, H.: Ber. **58**, 2175 (1925).
(26) Manchot, W., and Gall, H.: Ber. **59**, 1056 (1926).
(27) Manchot, W., and Gall, H.: Ber. **61**, 1135 (1928).
(28) Manchot, W., and König, J.: Ber. **57**, 2130 (1924).
(29) Manchot, W., and König, J.: Ber. **58**, 229 (1925).
(30) Manchot, W., and König, J.: Ber. **58**, 2173 (1925).
(31) Manchot, W., and König, J.: Ber. **59**, 883 (1926).
(32) Manchot, W., and König, J.: Ber. **60**, 2183 (1927).
(33) Manchot, W., and Manchot, W. J.: Z. anorg. allgem. Chem. **226**, 385 (1936).

[5] A recent review of the volatile metal carbonyls (2) contains a fairly extensive bibliography. In the present paper references contained in the earlier review are not always repeated.

(34) Manchot, W., and Pflaum, W.: Ber. **60,** 2180 (1927).
(35) Manchot, W., and Schmid, H.: Ber. **59,** 2360 (1926).
(36) Manchot, W., and Schmid, H.: Z. angew. allgem. Chem. **216,** 99 (1933).
(37) Manchot, W., and Waldmüller, A.: Ber. **59,** 2363 (1926).
(38) Manchot, W., and Woringer, P.: Ber. **45,** 2869 (1912); **46,** 3514 (1913).
(39) Pascal, Paul: *Traitè de Chimie Minerale,* Vol. IX, p. 544. Masson et Cie., Paris (1933).
(40) Reference 39, p. 840.
(41) Reference 39, pp. 858, 886.
(42) Reference 39, pp. 859, 887.
(43) Pascal, Paul: *Traitè de Chimie Minerale,* Vol. X, p. 207. Masson et Cie., Paris (1933).
(44) Reference 43, p. 409.
(45) Reference 43, p. 722.
(46) Reference 43, p. 961.
(47) Reference 43, p. 1071.
(48) Reference 43, p. 1089.
(49) Pascal, Paul: *Traitè de Chimie Minerale,* Vol. XI, p. 39. Masson et Cie., Paris (1932).
(50) Reference 49, p. 113.
(51) Reference 49, p. 365.
(52) Reference 49, p. 370.
(53) Reference 49, p. 376.
(54) Reference 49, p. 380.
(55) Reference 49, p. 505.
(56) Reference 49, p. 585.
(57) Reference 49, p. 669.
(58) Reference 49, p. 674.
(59) Reference 49, p. 675.
(60) Reference 49, p. 805.
(61) Reference 49, p. 809.
(62) Schützenberger: Compt. rend. **70,** 1134, 1287 (1870).
(63) Sidgwick, N. V., and Bailey, R. W.: Proc. Roy. Soc. (London) **144,** 521 (1934).
(64) Wilke-Dörfurt, E., and Balz, G.: Z. anorg. allgem. Chem. **159,** 219 (1927).

FACTS AND INTERPRETATIONS IN THE MECHANISM OF ALCOHOLIC FERMENTATION[1]

F. F. NORD

Department of Organic Chemistry, Fordham University, New York, New York

Received October 18, 1939

CONTENTS

"It is the great beauty of our science, that advancement in it, whether in a degree great or small, instead of exhausting the subject of research, opens the doors to further and more abundant knowledge, overflowing with beauty and utility."—FARADAY.

I. INTRODUCTION

The study of the biochemistry of alcoholic fermentation has displayed great variations in its development up to the present time. The changing knowledge applies not only to the research on the stoichiometric course of the reaction, but also to the fundamental considerations under which individual investigators attacked and pursued the complex question.

The following considerations of Turpin (168) were entertained already one hundred years ago:

[1] This article is an extension of a paper presented at the Third International Congress of Microbiology, held in New York, September, 1939. For the earlier status of the subject *cf.* F. F. Nord: Chem. Rev. **3**, 41 (1926).

The author gladly acknowledges the assistance of Rita C. O'Connor and John C. Wirth in the preparation of this paper.

"Que par fermentation on doit entendre: association composée d'eau, de corps vivants se nourissant et se développant par absorption, de l'une des parties du sucre, et en isolant, soit l'alcool, soit l'acide acétique; action toute physiologique qui commence et finit avec l'existence des infusoires végétaux ou animaux qui la déterminent, et dont la vie ne cesse que par l'épuisement totale de la matière saccharine et nutritive. C'est alors que mourants d'inanition et ne pouvant plus se soutenir dans l'épaisseur ou à la surface du liquide, on les voit se précipiter les uns sur les autres et s'entasser au fond du vase sous forme de lie mucilagineuse, de sédiment ou de levure."

The discussion to be presented does not require that a comprehensive account of the historical evolution of this fascinating subject be given. We shall, however, attempt neither to omit nor to diminish such difficulties as have resulted to a certain degree from the practical application of cell-free fermentation, discovered by Buchner (15) at the turn of the century.

Alcoholic fermentation by the action of undamaged cells has been studied during recent years by use of the enzyme systems of: (*a*) common top and bottom yeasts, (*b*) various fusaria, (*c*) *Thermobacterium mobile* Lindner, and (*d*) *Zymosarcina ventriculi*. Of these systems, the action range of the last two mentioned has been only "felt out." We know, however, that the thermobacterium (69, 79, 110) forms from sugar equimolecular quantities of carbon dioxide (about 45 per cent) and alcohol (about 41 per cent), also some lactic acid (about 6 per cent); and further, that the last mentioned fermentation (154), discovered by Goodsir about one hundred years ago, is predominantly of alcoholic nature, producing small quantities of hydrogen (0.5 per cent), acetic acid (6.6 per cent), acetylmethylcarbinol (1.7 per cent), and formic acid (0.8 per cent). The hydrogen comprises about one-fifth of the total volume of carbon dioxide evolved. These latter systems, as well as the yeast *Torulopsis pulcherrima* (136), will not be considered in the subsequent presentation.

The first enlightenment as to the course of the enzymatic breakdown of carbohydrates comes from two statements, valid until now: In the course of his important metabolism researches, Magnus-Levy in 1902 gave expression to the opinion that acetaldehyde was formed as a breakdown product of carbohydrate, and in 1910 O. Neubauer (105), in the course of his distinguished work on the fermentation of amino acids, announced this pioneer discovery:

"Weiter ist zu schliessen, dass die hier als Zwischenprodukt auftretende *Brenztraubensäure* durch gärende Hefe unter Reduktion *zu Kohlensäure und Alkohol* zersetzt wird, d. h. mit anderen Worten, dass sie leicht *vergärbar* sein muss. Eigens angestellte Versuche, die noch nicht völlig abgeschlossen sind, haben die Richtigkeit dieses Schlusses bestätigt. Damit ist nun ohne weiteres der Gedanke gegeben, die

Brenztraubensäure könnte ein *Zwischenprodukt bei der alkoholischen Gärung* des Zuckers sein; . . . ."

As a matter of fact, Fernbach and Schoen (37) in 1913 identified pyruvic acid in the presence of calcium carbonate in fermentation mashes containing living yeast cells, and isolated it in pure form (as the calcium salt) in quantity equal to 1.23 per cent of the original sugar.

About eight years later Neuberg, together with von Grab (47), confirmed this after he had in the first place rejected it, just as he had rejected the fermentation of pyruvic acid itself. Consequently, in 1919 he induced Kerb (66) to publish the following statement: "Brenztraubensäure, die schon auf Grund des Alkoholertrages nicht in nennenswerter Menge gebildet sein konnte, war auch *nicht* spurenweise qualitativ *nachzuweisen.*"

Because of Neubauer's discovery, the year 1910 is the experimental turning point from which our present knowledge on the course of enzymatic carbohydrate breakdown has developed.

II. CONCEPTS OF THE PHASE SEQUENCE OF ALCOHOLIC FERMENTATION

In alcoholic fermentation nature shows us a process which, in the case of the action of yeast, because of its optional anaerobic mode of life, will not lead to the oxidation end products, but, in contrast to alcoholic fermentation by fusaria (see later), ends with the formation of alcohol, an incompletely degraded product.

It is therefore not surprising that until now, about one hundred and thirty years after Gay-Lussac (42) set forth his basic equation

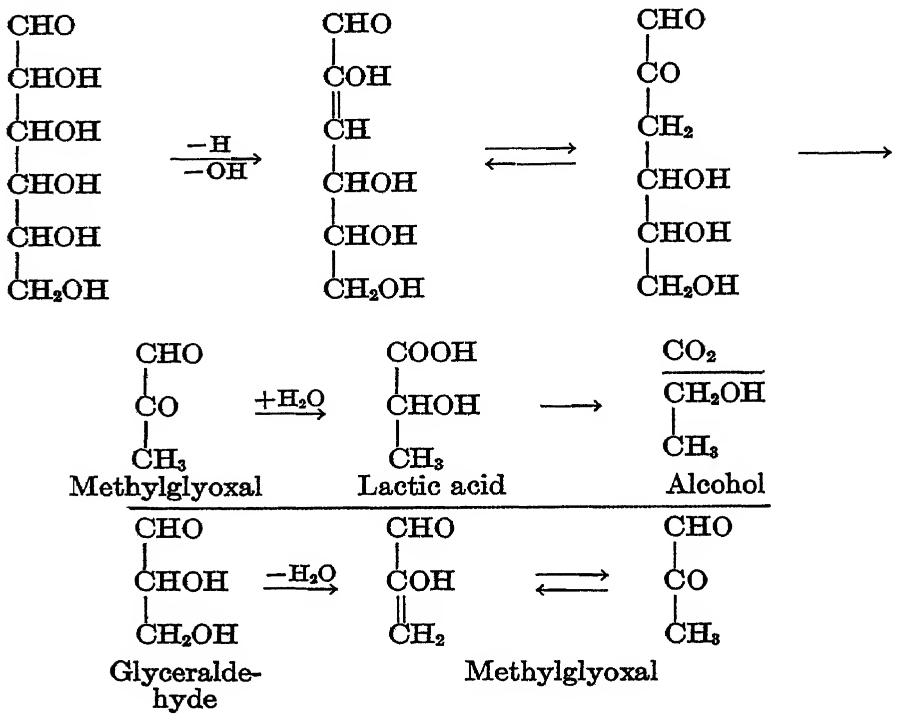
$$C_6H_{12}O_6 = 2CO_2 + 2C_2H_5OH$$

almost every natural scientist who occupied himself with studies of fermentation believed that he must establish his own theory as to the reaction mechanism.

The rapid development, which received its impetus from Neubauer's discovery, exempts us, however, from discussing individually in this review the many visions put on paper. They were more or less offsprings from the circle of thought of the great epoch of preparative chemistry. They were in ignorance of much which has become for us common knowledge of the experimentally proven mechanisms of enzyme action, and intimated that the breakdown product of sugar must be a smaller molecule than the fermentation substrate and must be easily split by the enzyme system of yeast.

The Wohl–Neubauer–Neuberg–Kerb reaction scheme has longest

withstood the stormy developments. Its central thought is based on the fact that a molecule of water can easily be split from an hydroxyl compound.



The six-carbon compound which will undergo further decomposition is, in this reaction scheme, methylglyoxal glyceraldehyde aldol, which, owing to easy hydrolysis, gives rise to two three-carbon split products, namely, methylglyoxal and glyceraldehyde. This possibility was considered in view of the findings of Pinkus (135) (production of methylglyoxal by heating glucose in alkaline medium) and of Wohl (181) (formation of methylglyoxal from glyceraldehyde itself).

On the basis of Neubauer's observation and the assumption of Magnus-Levy, Neuberg and Kerb (109) set up a remodelled reaction scheme in which the hypothetical methylglyoxal glyceraldehyde aldol of the Wohl scheme was replaced by the methylglyoxal aldol.

$$C_6H_{12}O_6 - 2H_2O \rightarrow \underset{\text{Methylglyoxal aldol}}{C_6H_8O_4}$$

$$C_6H_8O_4 = 2CH_2{=}C(OH)CHO \text{ or } \underset{\text{Methylglyoxal}}{2CH_3COCHO}$$

$$\begin{array}{l} CH_2{=}C(OH)CHO + H_2O \\ CH_2{=}C(OH)CHO \end{array} + \begin{array}{c} H_2 \\ \| \\ O \end{array} = \begin{array}{l} CH_2(OH)CHOHCH_2OH \text{ (glycerol)} \\ CH_3COCOOH \text{ (pyruvic acid)} \end{array}$$

$$CH_3COCOOH = CO_2 + \underset{\text{(acetaldehyde)}}{CH_3CHO}$$

$$\begin{array}{l} CH_3COCHO \\ CH_3CHO \end{array} + \begin{array}{c} O \\ \| \\ H_2 \end{array} = \begin{array}{l} CH_3COCOOH \text{ (pyruvic acid)} \\ CH_3CH_2OH \text{ (ethyl alcohol)} \end{array}$$

The presupposition of this sequence of phases was that the methylglyoxal would belong to that group of compounds which, in the course of the rearrangements taking place on the sugar molecule, are capable of taking over the rôle of a structurally intelligible *primary* biological split product. In addition, it was supposed to serve as one of the two different aldehydes which may undergo the mixed Cannizzaro reaction. In the course of this oxidation–reduction[2] the methylglyoxal should lead to the formation of pyruvic acid which, under the influence of carboxylase, is split into acetaldehyde and carbon dioxide. The acetaldehyde would then be reduced to ethyl alcohol and at the same time the methylglyoxal would be oxidized to pyruvic acid, thus continuing the cycle:

$$\begin{array}{c} \text{sugar} \rightarrow \text{methylglyoxal} \rightarrow \text{pyruvic acid} \rightarrow \text{acetaldehyde} + \text{carbon dioxide} \\ \text{methylglyoxal} + \text{acetaldehyde} \rightarrow \text{pyruvic acid} + \text{alcohol} \end{array}$$

The discovery of phosphoglyceric acid (113) among the products of the action of yeast juices or muscle extracts on carbohydrate, to which Embden (30) was the first to attach significance, led to observations and considerations which seemed at the time appropriate to bring nearer an understanding of the later phases of sugar dissimilation in the living cell as well. The essential feature of the new scheme was the fact that the methylglyoxal disappears therefrom and that the pyruvic acid, which, as before, without doubt represents an intermediate product in sugar decomposition, is formed from the phosphoglyceric acid.

$$CH_2O(PO_3H_2)CHOHCOOH \rightarrow CH_3COCOOH + H_3PO_4$$

As the next step, we see that the pyruvic acid is split by the action of carboxylase into acetaldehyde and carbon dioxide:

$$CH_3COCOOH \rightarrow CH_3CHO + CO_2$$

[2] Lavoisier, on page 101 of his *Traitè èlèmentaire de chimie* (Paris, 1864), wrote: "Les effet de la fermentation vineuse se réduisent donc à séparer en deux portions le sucre, qui est un oxyde, à oxygéner l'une aux dépens de l'autre pour former l'acide carbonique a désoxygener l'autre en faveur de la première pour enformer une substance combustible qui est l'alcool; en sorte que, s'il était possible de recombiner ces deux substances, l'alcool et l'acide carbonique, on reformerait du sucre."

In the phase of oxidation–reduction a further change, already indicated above, enters the scheme. As dismutation partner of the acetaldehyde, there enters in place of the methylglyoxal a triosephosphoric acid which has been synthesized by Fischer and Baer (40; *cf.* also 98) and the *d*-form of which is the readily fermentable (155; *cf.* also 171) 3-glyceraldehydemonophosphoric acid.

The oxidation–reduction phase may be represented as follows:

(*a*) $CH_2O(PO_3H_2)CHOHCHO \xrightarrow{\text{oxidation}} CH_2O(PO_3H_2)CHOHCOOH$
Phosphoglyceraldehyde — Phosphoglyceric acid

(*b*) $CH_3CHO \xrightarrow{\text{reduction}} CH_3CH_2OH$
Acetaldehyde — Alcohol

and furnishes the aforementioned phosphoglyceric acid, the mother substance of pyruvic acid. The cycle then will continue as long as the Fischer–Baer phosphoglyceric acid is present. It can be imagined that it is formed by the reaction:

$$1 \text{ glucose} + 2 \text{ phosphate} \rightarrow 2 \text{ glyceraldehydephosphoric acid}$$

The reaction once started continues as long as sugar and phosphate are present,—i.e., formation of glyceraldehyde, its oxidation to phosphoglyceric acid (with simultaneous reduction of acetaldehyde to alcohol), transformation of the latter to pyruvic acid, and finally to acetaldehyde and carbon dioxide,—and is ended by reduction of aldehyde to alcohol.

Apparently, the fulfillment of a presupposition is indispensable,—a trace of hexose diphosphate must be present before the reaction can begin. The hexose diphosphate is, however, not consumed, but acts as a catalyst. This has been referred to as the "stationary" condition. This phase of the reaction scheme was supposed to be initiated by the "Angärung," the onset of fermentation, in the course of which the required concentrations of reacting substances are built up. Here, again, the hexose diphosphate is required (and consumed), so that one molecule of it and one molecule of glucose plus two molecules of phosphoric acid form four molecules of phosphoglyceraldehyde:

$$1 \text{ hexose diphosphate} + 1 \text{ glucose} + 2 \text{ phosphoric acid} \rightarrow 4 \text{ phosphoglyceraldehyde} \rightarrow 2 \text{ phosphoglyceric acid} + 2 \text{ glycerophosphoric acid}$$

Since in this phase there is still no acetaldehyde formed which could take part in the oxidation–reduction of phosphoglyceraldehyde, it must be replaced by a second molecule of phosphoglyceraldehyde, of which one molecule is oxidized to phosphoglyceric acid and the other reduced to

glycerophosphoric acid, thus resembling the well-known transformation of acetaldehyde into acetic acid and ethyl alcohol.

$$2CH_3CHO + H_2O \rightarrow CH_3COOH + C_2H_5OH$$

$$\underset{\text{Phosphoglyceraldehyde}}{2CH_2O(PO_3H_2)CHOHCHO} + H_2O =$$

$$\underset{\text{Phosphoglyceric acid}}{CH_2O(PO_3H_2)CHOHCOOH} + \underset{\text{Glycerophosphoric acid}}{CH_2O(PO_3H_2)CHOHCH_2OH}$$

As soon as sufficient quantities of acetaldehyde are available, the reduction of phosphoglyceraldehyde slows down and the concentration of glycerophosphoric acid becomes slight. This was said to account for the appearance of slight traces of glycerol in the mash.

One of the various phase sequences which was formulated on this basis by Meyerhof (94),—but in the meantime also discarded,—is set up by the following reactions:

Onset of fermentation:

(*a*) 1 hexosediphosphoric acid + 1 glucose + 2 phosphoric acid = 4 triose-phosphoric acid = 2 glycerophosphoric acid + 2 phosphoglyceric acid

(*b*) 2 phosphoglyceric acid = 2 pyruvic acid + 2 phosphoric acid = 2 acetaldehyde + 2 CO_2 + 2 phosphoric acid

"Stationary" condition:

(*c*) 1 glucose + 2 acetaldehyde + 2 phosphoric acid = 2 triosephosphoric acid + 2 acetaldehyde = 2 phosphoglyceric acid + 2 alcohol

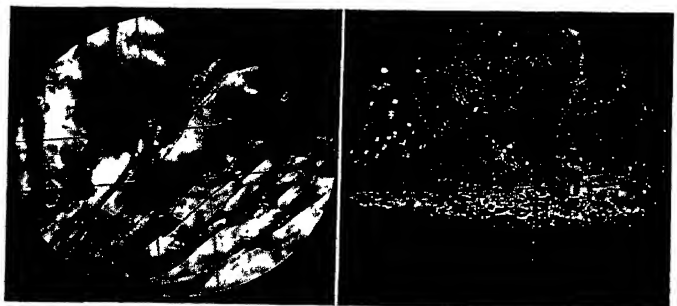
The greater part of the experiments serving as a basis for this scheme were conducted with the aid of juices in the presence of sodium fluoride. This inhibits phase *b* of the sequence, whereby the course is reduced to phase *a* in which a mixture of glycerophosphoric acids is formed. This acid appeared to Lohmann to be a difficultly hydrolyzable hexosephosphoric acid. Sodium fluoride does not inhibit the reduction of acetaldehyde or the oxidation of phosphoglyceraldehyde, and, in the presence of sodium fluoride, added acetaldehyde is further reduced, or, from glucose and phosphoric acid (in the presence of hexosediphosphoric acid) phosphoglyceric acid is produced as long as acetaldehyde is present.

III. BEHAVIOR AND PROPERTIES OF YEAST JUICES

It can be taken for granted that the methods for obtaining the different press and maceration juices are known. Until recently they were obtained chiefly from bottom yeasts. They form lyophilic colloidal dispersions,

which even after a storage period of 2 months at about $-5°C$. not only maintain their original activity, but show an increased rate of fermentation at the beginning of the proper experiments (124).

Comprehensive experiments have been conducted in an attempt to clarify this phenomenon. By the use of different lyophilic model-colloids the information was obtained that in different concentrations, under the influence of frost, the surface of the colloidal particles is either increased or diminished, and that the colloids are changed to such an extent that, among other things, they show measurable differences when exposed to the adsorption of various gases (120). It was this, among other evidence, which was seen as proof that freezing causes a disaggregation-aggregation of the particles. As a consequence of this change, exerted by physical means, it should be possible to bring about and observe an alteration in the



(a) (b)

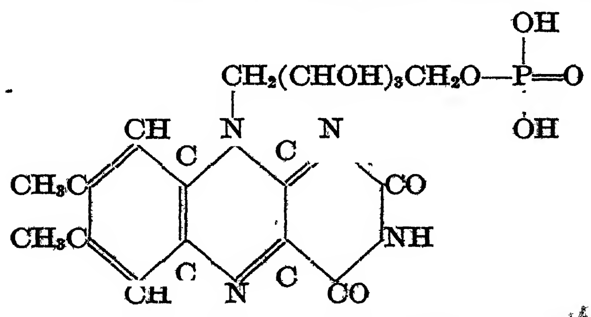
FIG. 1. (a) Disaggregation of a 1 per cent solution of egg albumin. Magnification 90 ×. Photographed at $-14°C$. (b) Aggregation of a 2 per cent solution of egg albumin. Magnification 90 ×. Photographed at $-10°C$.

particle radius. This change also appears, independently of the shape of the particle, in a difference in the rate of diffusion. Under the influence of frost, therefore, substances close to biological concentrations, in general, undergo a breaking up of the particles with simultaneous increase of their surface; at higher concentrations an increase in size, possibly to the point of coagulation, occurs with simultaneous decrease in surface area (see figure 1). Therewith the circle of proof is closed with regard to the later-mentioned dualistic or carrier theory of enzyme action, insofar as this applies to the colloidal-chemically acting part of the zymase system (117) and not to the structural-chemically acting part governing the manner and intensity of action as derived from the variability in the degree of dispersity.

The mechanism of the phenomenon also was cleared up recently (59, 60). If the activity of certain enzymes is explained by the carrier theory of

the mechanism of enzyme action through the common action of carrier and active group, then one must assume, in view of the aforementioned experimental findings, that in the case of a disaggregation a large colloidal particle breaks into two or more smaller particles, each of which, under all circumstances, is composed of carrier and active group. In other words, the crushed original colloidal particle is composed of several similarly constructed particles. A like conclusion follows from the aggregation process and, in connection with that, from the preceding experimental findings with so-called "protein-model bodies." Confirmation[3] of these experiments has been found also in the case of peroxidase, tyrosinase (41), catalase (76), certain viruses (145), and fibrinogen (140).

Alongside this colloid-chemical proof for the carrier theory of enzyme action we find flavin ferment (containing no polysaccharide) as the first enzyme to be observed from which more could be learned about the linkage between carrier and active group, or the purely physicochemical mechanism of its activity. The total enzyme, the elementary analysis of which is given by Theorell (165) as 51.5 per cent carbon, 7.37 per cent hydrogen, 15.9 per cent nitrogen, 0.043 per cent phosphorus, and 1.0 per cent sulfur, consisting of the carrier protein with flavinphosphoric acid as the active group,[4] can be obtained from Lebedew juice. The crystalline active group (165) of this "yellow ferment"[5] is identical with a derivative of vitamin B_2, lactoflavinphosphoric acid, which has the constitution of a 6,7-dimethyl-9-*d*-ribitylisoalloxazine-5′-phosphoric acid,



The accompanying, natural carrier protein, fully capable of coupling, can be obtained by dialysis of a salt-free aqueous solution of the pure enzyme,

[3] *Cf.* also H. B. Bull: Z. physik. Chem. A161, 192 (1932).

[4] Regarding its synthesis see R. Kuhn, H. Rudy, and F. Weygand: Ber. 69, 2034 (1936).

[5] The method of Theorell (Biochem. Z. 278, 236 (1935)) was meanwhile replaced by a simpler process of Warburg and Christian (Biochem. Z. 298, 368 (1938)): namely, splitting by action of hydrochloric acid in ammonium sulfate solution. Protein and prosthetic group can be obtained in good yield within an hour. Compare also P. Karrer: Ergeb. Vitamin- und Hormonforsch. 2, 381

whereby the linkage between the flavinphosphoric acid and the protein may be reversibly broken. The carrier protein, of albumin character, has approximately the same isoelectric point and rate of electrophoretic migration as the complete enzyme. From measurement of the reaction velocity it may be concluded that the "yellow ferment" can neither react in a noteworthy measure with molecular oxygen nor, in the physiological cell metabolism, with oxidized cytochrome. Under these circumstances it can be doubted that the flavin enzyme is a means for the oxidation of dihydrocozymase in the cell. This might be the task of the newly discovered diaphorase (1, 26), the realm of action of which has still to be marked off in relation to the neoflavin (49) enzyme. It has recently been found to be identical with the flavoprotein of heart muscle (23).

The findings elucidated here represent the experimental proof of the concept first put in words by Mathews and Glenn (91) as to the dual theory of enzyme action:

> "What we ordinarily call an enzyme, such as invertase, diastase, pepsin, etc., is a combination of a colloid with an active principle. The active principle is the enzyme itself and should of course be called the enzyme, but it has happened that the substances isolated as enzymes have been generally the combination of this active principle with the inert substance, colloidal in nature.
>
> "The colloidal part of the molecule which is inert might with propriety be called the zymophore or ferment bearer, since in cells most of the enzymes are probably thus united or borne, but as this word has been used by Ehrlich to designate the active principle itself, we may call the colloid simply the carrier or bearer, and the active principle, the enzyme or kinase."

With regard to the colloid-chemical part of the proof, (see page 430) that function of the carrier must also be taken into account which, by changing the electronic configuration of the substrate by activation, can bring about a corresponding exchange of electrons with the pyridine system (see later).[5a]

It is, however, time to call attention to the fact that in the further course of the discussion, a connection between mode of action and active group (for instance, of cocarboxylase) comes to the fore, which leaves doubts concerning its presence originally as an active group of specific proteins in the living cell.

Contrary to the general assumption that for the preparation of active maceration juice according to Lebedew (75) only dry preparations of bottom brewers' yeast are suitable, F. Lipmann (83) was able to show that from bakers' yeast, after suitable drying and upon addition of phosphate, a highly active juice can be obtained by extraction. The fermentation by bakers' yeast juice differs greatly from that by bottom yeast

[5a] See also reference 171, especially p. 46.

juice. As long as there is present a great excess of inorganic phosphate, the fermentation follows the Harden-Young equation, where each mole of carbon dioxide formed corresponds to one mole of phosphate esterified.

If, however, the phosphate content of the bakers' yeast extract sinks to about 20 per cent of the original concentration, the fermentation now proceeds at only a slightly lower rate without simultaneous disappearance of the remaining phosphate, while with bottom-yeast extract the fermentation, after consumption of free inorganic phosphate, falls abruptly to a very slow rate. Therefore the fermentation here proceeds from an esterification phase to a phase without esterification, in which not the Harden–Young but the Gay–Lussac equation holds. In this case, the cell-free fermentation proceeds during this phase like fermentation in the living cell. Very noteworthy was the observation made, in the course of these experiments, of the significant rise in the CO_2/P quotient, i.e., the increasing recession of esterification in relation to the fermentation. The bakers' yeast extract has a yellowish color, which is largely due to the content of the "yellow ferment."

It seems suitable for the purpose to refer to this difference in the course of fermentation by the two closely related extracts in order to be able to emphasize the significance of the later constant fermentation rate, since, according to the recent reinvestigations (115) of older findings, it is known that with all juices studied hitherto, the "further" fermentation has already reached a diminished rate before esterification has ceased.

The experimental background of these observations is to be recognized in results first described by Wroblewski (184). He found that sodium phosphate increased the activity of pressed juice. Moreover, Iwanow (63) showed, in 1905, that living yeast, like other plants, transformed inorganic phosphate to organic phosphate.

Harden and Young (53), in the course of their researches, then observed that the rate of fermentation of a press juice is rapidly increased if, to the fermenting mash, in the presence of soluble phosphate, boiled extract is added. At the end of the fermentation no appreciable amounts of phosphate could be detected by the usual means. The investigators concluded from their findings that two sugar molecules are involved in fermentation. While one sugar molecule, together with two moles of phosphate, forms hexose diphosphate,[6] another molecule forms alcohol and carbon dioxide. In the further course of the reaction the inorganic phosphate can be regenerated by the action of phosphatase on hexose diphosphate, and the fermentable hexose set free. Thus the cycle can begin

[6] The hexose diphosphate when heated with oxalic acid in aqueous solution furnished a mixed ester which consisted, however, predominantly of a keto- or monoester, called also Neuberg ester (Biochem. Z. 88, 432 (1918)).

again. By kinetic measurements Harden and Young were, moreover, able to determine that the quantities of carbon dioxide and ethyl alcohol evolved were (within certain limits) proportional to the added phosphate.

The equations set forth by them read as follows:

$$2C_6H_{12}O_6 + 2M_2{}^*HPO_4 = 2CO_2 + 2C_2H_5OH + 2H_2O + C_6H_{10}O_4(PO_4M_2)_2$$

$$C_6H_{10}O_4(PO_4M_2)_2 + 2H_2O = C_6H_{12}O_6 + 2M_2HPO_4$$

*M = metal.

The fact that Harden and Robison (50) had found, besides the diphosphate, also a mixed monophosphate, later closely investigated by Robison (147) (the Robison ester) led Raymond (142) to interpret the original equations of Harden and Young in a different way.[7]

The true rôle of phosphates in the metabolism of living yeast is even now not clarified beyond doubt. It may, however, be regarded as certain that, even if yeast cells do break down carbohydrates by a detour of intermediary phosphorylation, this is not necessarily the only way in which the degradation is accomplished.

In contrast to Buchner's cell extract or the often-studied maceration juices, the frozen extract, first obtained by Dixon and Atkins (27), represents a system which, above all, has the advantage that it also contains the insoluble enzymes otherwise left behind in the cell fragments. Moreover, its preparation can be accomplished with no autolysis whatsoever and the extract can be evaporated to dryness without change in its enzymatic potentialities.

The first investigators to study this juice more closely were Tait and

[7] *Cf.* F. F. Nord: Chem. Rev. **3**, 50 (1926). During the preparation of this paper there appeared a short preliminary communication by E. Negelein and H. Brömel (Biochem. Z. **301**, 135 (1939)), according to which the diphosphoglyceric acid (see page 427) isolated by them from Lebedew juice is removed by the action of various enzymes and coenzymes, whereby the total phosphate is transferred to the sugar. In the living yeast cell, half of the phosphate from the above acid is supposed to be transferred to the sugar and the other half is supposed to be freed:

$$\text{phosphoric acid} + \text{glyceraldehydephosphoric acid} \rightleftharpoons \text{glyceraldehydediphosphoric acid}$$

In the presence of diphosphopyridine nucleotide and a specific carrier protein, the following sequence is supposed to occur:

$$\text{pyridine} + \text{glyceraldehydediphosphoric acid} \rightleftharpoons \text{dihydropyridine} + \text{glyceric acid diphosphoric acid}$$

At first glance, however, the presentation lacks the explanation required for the understanding of the mechanism of the indispensable intermediary reactions in the living cell. Concerning the properties of the acid, consult Biochem. Z. **303**, 132 (1939).

Fletcher (162), who determined the hydrogen-ion concentration of the frozen extract to be 6.2. This specification is in accord with the results

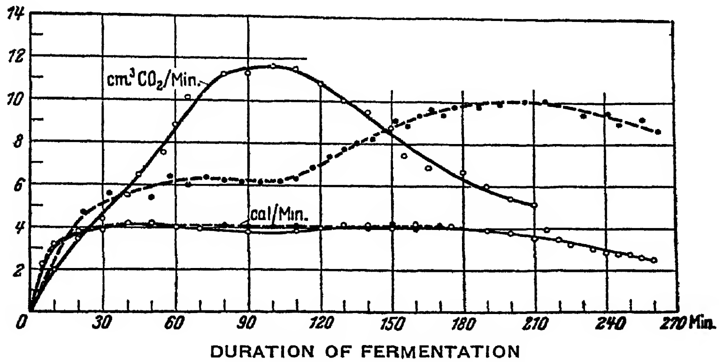


FIG. 2. Course of reaction using fresh living yeast (2 g. in 200 cc. of a 4.5 per cent solution of glucose). Experiment No. 1: ----- carbon dioxide evolved; —— heat evolved. Experiment No. 2: —— carbon dioxide evolved; ----- heat evolved.

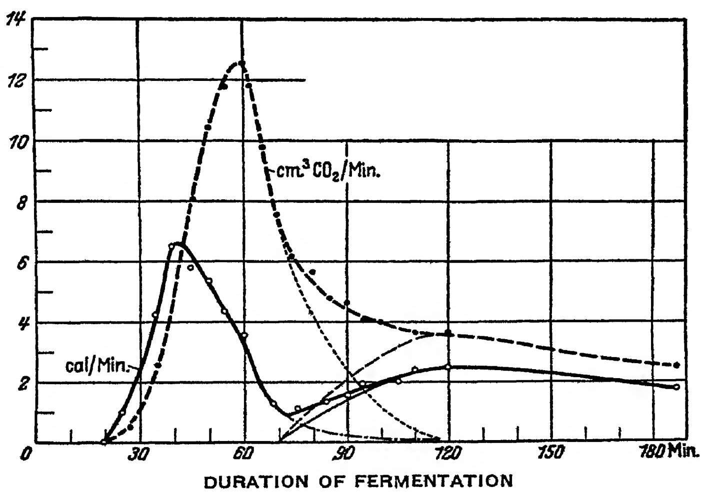


FIG. 3. Course of reaction using maceration juice (9 g. of glucose in 200 cc. of juice). -·-·-·- extrapolation of the fermentation curves of free sugar; ---- extrapolation of the fermentation curves of the products of phosphorylation.

of measurements by Mahdihassan (89) of the internal hydrogen-ion concentration of different living-cell systems.

Each finding which concerns the similarity or dissimilarity of the course of material or energetic conversions in the living cell, or by action of

enzyme preparations, is of importance for the understanding of the mechanism of enzyme action or the course of the phase sequence. The function of the cell consists in the liberation of energy. Energy given off in the form of heat is lost. Therefore the relation of two reaction products of alcoholic fermentation,—the carbon dioxide evolved and the energy liberated as heat,—were compared during fermentation by living yeast cells and by Lebedew extract (58) (see figures 2 and 3). It was thus noted that the heat of reaction in the course of fermentation changed continually. This meant that fermentation with living yeast does not proceed according to a fixed outlined scheme. The thermochemical course of fermentation with yeast maceration juice shows, in contrast, that here at least two different conversions occur: (*a*) the fermentation of free sugar in the presence of free phosphate (inhibited by phloridzin); and (*b*) the subsequent fermentation of the residual substrate in the absence of free phosphate (not inhibited by phloridzin). Neither reaction shows a resemblance in thermochemical respects to fermentation with living yeast.

The phosphorylation of sugar in maceration juice is accompanied by a heat of reaction of the order of magnitude of 22 cal. per millimole of phosphate esterified. This value corresponds to almost the total heat which is evolved during the decomposition, sugar → alcohol + carbon dioxide: namely, found, 24 cal.; calculated, 28 cal.[8] In general, the relation between the course of the actual reactions in the cell and those found individually in destroyed systems and united afterwards in a scheme may be considered as being the same as the relation between a quotient of differences and a differential quotient.

IV. DISCOVERY OF ZYMASE AND COZYMASE

The (previously mentioned) discovery of cell-free fermentation by Buchner and Hahn had far-reaching consequences. In the dispute between Liebig and Pasteur concerning the recognition of the essence of alcoholic fermentation, it signified that neither of the last named was right but that, at the time, a third came nearest to the truth. Traube had postulated, as early as 1858, that all of the fermentations brought about by living organisms were caused by enzymes secreted by the cells.

Buchner and Hahn first used their extracts in animal studies and noted that they changed considerably in a short time. To protect them against deterioration and loss of activity, they tried, without success, the usual chemically acting preservatives and therefore added sugar to them. This experiment became the first step in the study of cell-free fermentation. In the work mentioned Buchner was able to determine that the

[8] Regarding the heat of decomposition for other conversions in the sphere of carbohydrate breakdown, compare L. Genevois: Ann. fermentations 2, 65 (1936).

cell-free, thermolabile yeast juice induced the fermentation of various monosaccharides and of maltose, and that this capability was not lost either by the action of chloroform, benzene, or sodium arsenate or by filtration, evaporation, or precipitation with alcohol.

From these facts, E. Buchner arrived at the following fundamental conclusions:

> "Zunächst ist bewiesen, dass es zur Einleitung des Gärungsvorganges keines so complizierten Apparates bedarf, wie ihn die Hefezelle vorstellt. Als Träger der Gärwirkung des Pressaftes ist vielmehr eine gelóste Substanz, zweifels ohne ein Eiweisskorper zu betrachten; derselbe soll als Zymase bezeichnet werden."

If the extract thus obtained was filtered through a Chamberland candle, the subsequent fractions displayed gradually diminishing powers of fermentation until no activity at all was shown. The juice contained not only zymases but also digestive enzymes which split proteins down to the amino acids. Hydrocyanic acid, in contrast to its action on fusaria (see later), produced a complete but reversible suppression of the juice fermentation.

By addition of boiled yeast juice,—the so-called "Kochsaft,"—alcoholic fermentation, which is caused by a certain volume of pressed juice, is considerably increased. This observation was the starting point for another great discovery. Harden and Young (52) separated top-yeast juice into an inactive filtrate and an inactive residue by means of ultrafiltration (Martin). Buchner and Antoni (16) had obtained the same results at an earlier date by the dialysis of yeast juice. When the filtrate and residue, neither of which alone could produce fermentation, were again united, an alcoholic fermentation of sugar took place. The residue, filtered, washed, and freed of inorganic phosphate, consisted chiefly of glycogen, dextrins, and proteins, as well as other enzymes present in the juice. It would not ferment sugar. Harden and Young concluded from these experiments that the fermentation of sugar induced by yeast juice is dependent, aside from zymase, on the presence of a dialyzable system which, according to Tholin (166), is thermostable to about 80°C. According to a reference of Bertrand (13) they called the system coenzyme. von Euler and Myrbäck (32) proposed its present name, cozymase (see later).

Neuberg and von Euler (107) have specified the following nomenclature: By zymase is meant the total enzyme system involved in alcoholic fermentation free of activators. By holozymase is meant the enzyme system of fermentation consisting of zymase plus all the activators. By apozymase is meant the holozymase free of cozymase.

Some of the investigations of von Euler and Myrbäck (32) may be considered as preliminary studies toward the isolation of cozymase and the detection of its manner of action (see page 460).

V. KINETICS OF CELL FERMENTATION

If one is trying to investigate the physical and chemical suppositions which are suitable to account for the realization of a cell reaction, then it is obviously necessary, in order to simplify conditions, to use unicellular systems and to start from the following considerations: In order that a material change can take place in the cell it is necessary that the substrate pass into the reaction space, i.e., permeate through the cell membrane,

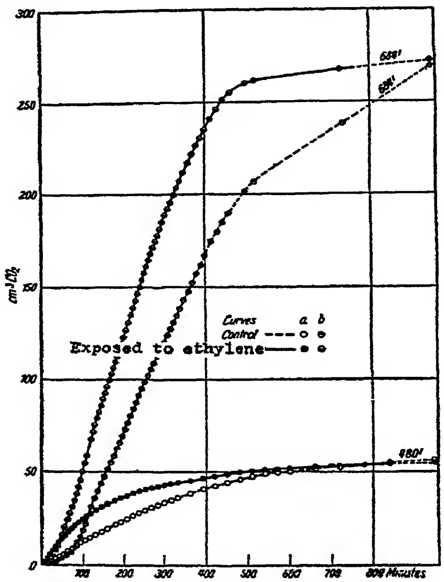


Fig. 4

and that, after entrance into the reaction space, it be chemically or enzymatically converted, whereby, in the latter case, the substrate molecule and catalyst must come into appropriate contact with one another.

In the course of extensive researches[9] on the influence of cell perme-

[9] See F. F. Nord: Z. angew. Chem. **42**, 1022 (1929); Science **79**, 159 (1934); Food Manuf. **9**, 55 (1934); Ergeb. Enzymforsch. **1**, 77 (1932). The proof that ethylene, acetylene, and related compounds can influence enzymatic reactions, on the one hand accelerating by increasing cell permeability, and on the other hand retarding through being reversibly taken up by the lyophilic colloid elements, is demonstrated in figure 4. In the illustration the lower pair of curves shows the course of carbon dioxide evolution in a fermentation by living top-yeast cells of a 2 per cent solution of pyruvic acid. The essentially steeper upper pair of curves illustrates the course of fermentation of sugar. The sugar fermentation was carried out with the same yeast which was used on the pyruvic acid, after being repeatedly washed and centrifuged.

ability[10] on enzymatic reaction, the possibility of renewed study of the kinetics (126) of fermentation by live yeast presented itself.

The earlier results are founded on the hypothesis that one is dealing here with a reaction course the speed of which is proportional to the quantity of yeast, and which can be represented as a first-order reaction. In contrast to that, earlier studies have shown that the rate of fermentation is approximately independent of the glucose concentration within the range of 0.5 per cent to 10.0 per cent. Insufficient notice was taken of this contradiction. Here we are dealing not with a reaction course which is influenced only by the amount of enzyme surface, but with a breakdown of the sugar molecules which must first pass through a membrane. The aforementioned contradiction is adjusted by this concept. It was tested by the influence of different speeds of stirring on the reaction course.[11] By exerting this influence on the fermentation action it was established that under the prevailing experimental conditions the sedimentation of cells, which was considered as responsible for the inhibition of diffusion in the outer medium, could be compensated by a rotation speed of twenty turns per minute. Up to one hundred fifty turns per minute produces a proportionate rise in the speed of fermentation (see figure 5). Since the differences in diffusion are already compensated at a stirring speed of twenty turns per minute, this further acceleration of the reaction is due solely to the relative membrane motion in regard to the substrate solution or to the increase in membrane surface, whereby more sugar is allowed to permeate into the cell. By running fermenta-

[10] According to the observations made by B. Luyet (Compt. rend. **204,** 1214, 1506 (1937); Compt. rend. soc. biol. **125,** 403 (1937)), the action of increased pressure or of heating is identical. Both destroy the permeability of yeast cells in the same way.

[11] The course of the rising branches of the curves submitted, especially the progressive shortening of the onset phase, and the observations on the behavior of a frozen zymase solution (compare also table 1 in the work of Nord and Franke (124) and Ergebnisse der Enzymforschung **1,** 79 (1932)) are in good agreement with the onset of the curves in figure 11 of an investigation by F. Lynen (Ann. **539,** 1 (1939)) on the behavior of frozen extracts. Lynen assumed that the progressive flattening of the upper parts of his fermentation curves could be traced back to the addition of increasing quantities of frozen yeast. We have, on the other hand, proven (*a*) that the initial ascent of the curves of fermentation and its consequences are governed by a physical effect and (*b*) that cryolysis in biological concentrations discloses a disaggregation of the carrier particles. In contradiction to the *conclusions* of Lynen, based on the chemistry of enzymes, there are to be found our *confirmed results* concerning the kinetics of cell fermentation and the increased activity caused by freezing of the carrier systems concerned. So far as the effect of freezing is concerned there is no question whether enzymes will be damaged through the action of low temperatures, since with the aid of cryolysis the proof of augmenting their activity was produced in the case of different enzymes.

tions with different initial concentrations of sugar, the conclusion was reached that the reaction showed no maximum rate at a concentration of 4 per cent—contrary to the findings of Slator (153)—but continued to rise beyond this point (see figure 6). At higher concentrations the per-

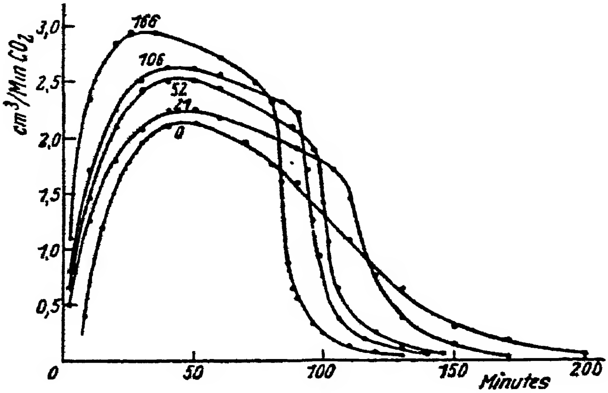


Fig. 5. Effect of stirring on the rate of fermentation

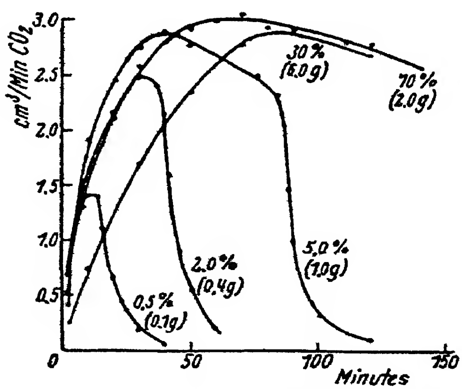


Fig. 6. Effect of concentration of sugar on the rate of fermentation

meability of the membrane is decreased. The graph showing the reaction course itself (time:rate) corresponds to the exponential function

$$-\frac{d\sigma}{dt} = \frac{\mu k}{k - \mu} S_0(e^{-kt} - e^{-\mu t})$$

in contrast to the former consideration which was represented by a straight line. It is the solution of the differential equation

$$\frac{d^2 c}{dt^2} + (\mu + k)\frac{dc}{dt} + \mu k c = 0$$

in which μ is a constant, c is the concentration of sugar in the cell, and k is the unimolecular reaction constant.

Similar experiments were later carried over to multiplying yeast cells. The aforementioned results were hereby confirmed (68, 61).

Medwedew (92), assuming that the energy of activation for alcoholic fermentation amounts to only 5000 cal., calculated, for a change of 500 cal. within a range of 10°, the high temperature coefficient of 3.0.

If one henceforth tries to study fermentation from the above explained and experimentally and theoretically founded standpoint, then it is possible to obtain fundamental knowledge and criteria for the interpretation of the biological and chemical reaction mechanism of fermentation. A difference worthy of notice is shown by comparison of the rate

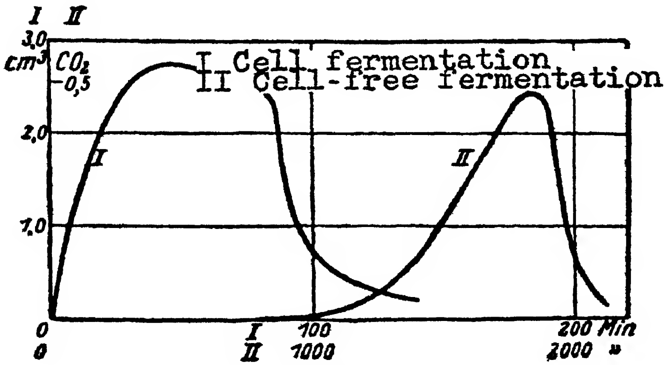


FIG. 7

curves of a fermentation with living cells which shows an acceleration at the onset of the reaction, and one which is brought about by cell-free juice (figure 7). Since the rate of substrate conversion on the enzyme surface, which doubtless is governed by a certain relation between the specificity of the carrier and the structure of the active group, or the molecular structure of the substance, is independent of the speed with which the mash is stirred, it has to be concluded from the course of figure 5 that the occurrence of enzymatic substrate conversions in the living cell is decisively influenced by the rate of diffusion. An effect of this kind on the reaction rate is impossible in a system of destroyed structure,[11a] and therefore draws attention to the importance of the assumption that the dissimilation of phosphorylated basic substances or their derivatives in the living yeast cell may at least be paralleled by the direct (phosphorus-

[11a] This is not contradicted by certain findings of Lipmann (Biochem. Z. **274**, 414 (1934)), because his observation of an accelerated induction in a maceration juice shows the influence of an oxygen atmosphere as compared to one of nitrogen.

free) breakdown of a substrate molecule, just as in an embryonal tissue (103)[12], rabbit erythrocytes (127, 186), propionic acid bacteria (172, 178), as also with living fusaria (see later), or during the oxidation occurring in the cerebral cells (64). By controlled pH relations (see page 449), the so-called intermediate *products* were accordingly (in the first case) regarded solely as *phases* of transformation without having definite molecular structure in the usual sense. The fragments were, in the course of degradation, stabilized to a certain degree by adsorptive power, and thus their dislocation was regulated.

It appears that the above conclusions would be weakened on the ground of experiments of von Hevesy, Parnas, Ostern, and others[13]. These investigators prepared radioactive adenylic acid from radioactive phosphate, $Na_2H^{32}_{15}PO_4$, and adenosine, and added it to fermenting yeast. After the lapse of appropriate incubation, they could recover a "significant part" of the radioactive phosphate bound in the sugar phosphoric acids.

However, one of the collaborators, von Hevesy (56) himself, gave the following interpretation of the results thus far obtained:

> "Although the investigation of the formation of labelled acid-soluble P compounds, both by *in vivo* and by *in vitro* experiments, supplies us with valuable information as to the formation of organic compounds of this type, the very appreciable speed with which some of them are resynthesised in the organism somewhat restricts the applicability of isotopic indicators in the study of their formation."

The expectations which were attached to the strength of evidence of the experiments carried out with the aid of ^{32}P, were consequently not fulfilled (99).

In accordance with our earlier statements, the conclusion is justified that induction bears no importance to cell fermentation. Consequently, there is no common ground, in this respect, to be found for a direct comparison of a fermentation caused by living cells with cell-free fermentation, where, even according to Meyerhof (95), the evolution of carbon dioxide is preceded in the same period by esterification.

[12] Compare, for example, the very foresighted statement of E. Grafe (Biochem. Zentr. **6**, 446 (1907)): "Diese Tatsache ist von höchstem Interesse, da in dieser Zeit der Grundstock für die Organe in Embryo gelegt ist. Es würde sich also dieser Zeitpunkt von dem an die Entwicklung allein durch Wachstum der nun differenzierten Gewebe vor sich geht, auch chemisch scharf . . . markieren." *Cf.* also G. von Hevesy, H. B. Levi, and O. H. Rebbe: Biochem. J. **32**, 2147 (1938).

[13] Quoted from P. Ostern, T. Baranowski, and J. Terszakowec (Z. physiol. Chem. **251**, 261 (1938)).

VI. IS PHOSPHORYLATION ESSENTIAL TO ALCOHOLIC SUGAR BREAKDOWN IN THE CELL?

At the very outset of the task of outlining the breakdown of hexoses by the enzymes of yeast systems, a fundamental difficulty presented itself. Aside from the fact that, as was previously mentioned, no possibility exists of becoming acquainted with a definite course which would characterize the phase sequence of degradation, the initial phase of substrate mobilization seems obscure. Even today, after the lapse of twenty-five years, the relation is best expressed by the following statement (148);

"Conceptions of the breakdown of glucose in which the process is depicted as though it proceeded from one fixed molecule to another, are fundamentally inadequate."

A starting point for the forthcoming explanation (which was noted late) is to be found in a contribution of Somogyi (157), who ascertained that when a yeast suspension and a glucose solution are appropriately brought together in the presence of protein, no sugar can be found in the filtrate. His assumption that here we dealt with adsorption was strengthened by the establishment of the specificity of the phenomenon for fermentable sugars as opposed to the case of lactose, arabinose, etc., which are not adsorbed on the yeast cell surface under the same conditions, but can be quantitatively recovered. One can compare this observation, in the spirit of present terminology, with the spreading, deformation, or loosening of the linkages of the substrate molecule by the adsorptive power of enzymes present in or on the cell. This introduces a molecular structure which is called the fermentable sugar form or "transport form" (119)[14] and which was thoroughly discussed, without interpreting it as a case of spatial isomerism in the classical sense of structural chemistry. On the basis of the above considerations (page 442), it is conceivable that a deformed molecule disintegrates in the cell into labile C_3 fragments without uniting with the phosphoric acid present, and essential for the building up of the body of yeast cells.

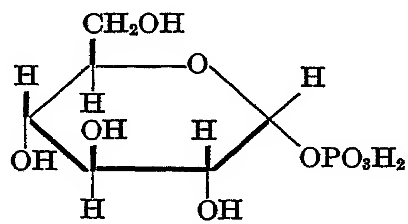
Willstätter and Rohdewald (180) have taken an opposite standpoint, after the prior work of Grüss (48), in support of Wertheimer (173), and they claim to have found 99 to 100 per cent of the sugar removed from solution in the form of a polyose. This value is for top yeast; about 50 per cent was claimed in the case of bottom yeast. The transformation of sugar to the glycogen complex, composed of glycogen, yeast gum, and a membrane polyose, "seems" to be followed by a glycogenolysis, which "probably" transforms it into a sugar form more capable of reaction

[14] Wieland later made use of an analogous concept in connection with considerations of the formation of succinic acid in the enzymatic oxidation of acetic acid (Helv. Chim. Acta **15**, 521 (1932)).

(see above), which is subject to phosphorylation. The observation of Somogyi was criticized by Benedict (11) and also by Raymond and Blanco (143), although the enzymatic attackability (139) of glycogen, under the experimental conditions prevailing and corresponding to the protein content of yeast, varied. According to Parnas (131), glycogen, and according to Ostern (128), starch[15] also, but not inulin, can be enzymatically split and transformed into hexose monophosphate according to the equation:

$$\text{glycogen} + \text{phosphate} = \text{hexose monophosphate}$$

This process was justifiably called phosphorolysis (instead of hydrolysis) by Parnas and is supposed to be reversible (149a; *cf.* also 31). But Meyerhof recently induced Goda (45)[16] to maintain that he had found "keine Anhaltspunkte fur die Annahme von Willstätter und Rohdewald, wonach in der lebenden Hefe der Zucker nur auf dem Wege der Glycogensynthese vergärbar wäre." This utterance is especially noteworthy in view of the discovery by Cori, Colowick, and Cori (20, 20a)[17] that glucopyranose-1-phosphate



is formed in washed muscle pulp in the presence of adenylic acid[18] by incubation of glycogen with inorganic phosphate (and magnesium) and was supposed to be observed also in yeast juice fermentations (150, 121).

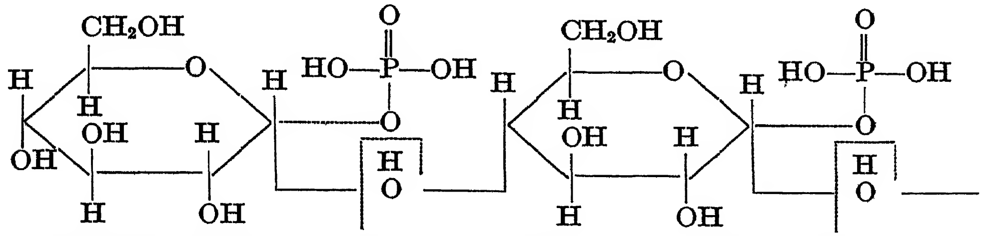
[15] According to Sarzana and Cacioppo (Biochim. terap. sper. **25**, 359 (1938)), injected starch is utilized by animal tissues without previous transformation into glucose or glycogen.

[16] Kruyk and Klingmüller (Biochem. Z. **300**, 343 (1939)) were also unable to show a synthesis of polysaccharide in yeasts of different ages. This is in agreement with recent findings of W. Kiessling (Biochem. Z. **302**, 70 (1939)), who also found that the primary step in the course of glucose fermentation does not involve formation of glycogen.

[17] Concerning the preparation of the crystalline potassium salt of the Cori ester, compare W. Kiessling: Biochem. Z. **298**, 421 (1938).

[18] In experiments undertaken by Bauer and von Euler (Z. physiol. Chem. **255**, 89 (1938)) serving to reëxamine the observations of Kendal and Stickland (Biochem. J. **32**, 572, (1938)), there was no clear stand taken. The latter were found to be in good agreement with the finding of Parnas or of Cori and Cori, showing that during the course of phosphorylation of glycogen through muscle preparation the action of adenosinetriphosphoric acid is much superior to the action of adenylic acid.

The glycogen molecule was thus split, according to Cori and Cori, without water taking part in the reaction.



The diagram shows the entrance of inorganic phosphate at the maltose linkage.

This is again in harmony with the above-mentioned glycogen formation as a primary step of fermentation in the living cell, whereas the reversibility of the phosphorolysis of glycogen and the alleged separability of the phosphorylating system of glucose or glycogen stands at variance with it. The conversion of the Cori ester is, however, not of a uniform nature (151). Besides the 6-ester, there originate also substances giving the same color reaction with iodine which is characteristic for glycogen. By contact with soluble muscle ferments this non-reducing Cori ester is transformed to glucopyranose-6-phosphate. The transformation is greatly accelerated by magnesium ions (19), proceeds instantaneously, and seems not to be an equilibrium reaction. From the glucopyranose-6-phosphate is formed, —by the action of phosphohexokinase (163) (oxoisomerase of Lohmann),— fructofuranose-6-phosphoric acid, the Embden ester (hexosemonophosphoric acid), which is an equilibrium mixture consisting of Robison and Neuberg esters in the ratio of 85 per cent aldose to 15 per cent ketose.

In this connection it is important to record the mobility of opinion concerning the interpretation of the author's own experiments. In 1936 A. Schäffner (149) wrote:

"Die katalytische Reaktionsweise des organisch gebundenen Phosphates lässt die weitgehende Schlussfolgerung über die Entbehrlichkeit des Phosphates bei der alkoholischen Gärung u. E. vorderhand als nicht berechtigt erscheinen."

In 1938, on the contrary, in the previously mentioned work (150) he expressed himself as follows on the question of the appearance of glycogen at the onset of fermentation:

"Die Vergärung von Glucose mittels zellfreien Hefeauszügen geht—mindestens zu einem grossen Teil—nicht über Glykogen. Die Feststellung, dass in zellfreien Extrakten ein anderer Weg zu Hexosemonophosphat führt als in der lebenden Zelle, rührt an das Problem, wie weit überhaupt von Ergebnissen, die mit zellfreiem Material erreicht werden, auf Vorgänge in der Zelle geschlossen werden darf. Wenn auch die Schlüsse mit aller Vorsicht gezogen werden müssen, so würde es u. E. zu

weit gehen, den Reaktionen, die in vitro festgestellt werden können, jeden biologischen Sinn abzusprechen. Es ist so gut wie ausgeschlossen, dass bei der Isolierung aus der lebenden Zelle ein ganz neuer Apparat von Enzymen auftritt. Viel wahrscheinlicher ist es, dass dabei ein Enzym verloren geht oder dass die Konzentrationsverhältnisse der Enzyme geändert werden. So darf man auch in unserem Fall annehmen, dass alle Phosphorylierungsmechanismen, die in vitro nachgewiesen werden konnten, auch in der lebenden Zelle eine Rolle zu spielen haben, dass aber entweder die Verkettung der Phosphorylierungsmechanismen eine besonders geartete ist, die wir nicht kennen und die in zellfreien Extrakten gestört ist, oder dass die lebende Zelle die verschiedenen ihr zur Verfügung stehenden Wege der Phosphorylierung einschlagen kann, je nach dem Bedürfnis und der Zweckmässigkeit."

Wieland, Claren, and Wille (175) have reported that bottom yeast undergoes no functional damage as a result of the aerobic removal of its nutrients. Such a yeast, which can be obtained by shaking an aqueous suspension for 15 to 20 hr. in an oxygen atmosphere, is, after this procedure, completely intact morphologically and physiologically. It is especially suited to the study of transformations having reaction rates which are insignificant as compared to the representative reactions of yeast which has not been pretreated. This "impoverished" yeast, in the conversion of alcohol, behaves in an entirely different manner from yeast which has not been pretreated, with which, according to Meyerhof, a carbohydrate resynthesis in the ratio: one part of alcohol → two to five parts of carbohydrate is supposed to occur. In the presence of the "impoverished" yeast, on the other hand, of fourteen parts of alcohol, eleven are oxidized, fat is formed from two parts, and only one part is converted into carbohydrate (177). Top yeast, when shaken under oxygen, gives up after 4 to 6 hr. the greater part of its dissimilable contents (176)[19]. So far as the author knows, no juices have as yet been prepared from this "impoverished" yeast. It is also necessary to determine how this yeast or the preparations obtained therefrom behave with reference to esterification. What would be the first phase of dissimilation in the case of a yeast which is fermentatively active but free from polyoses?

An investigation of this question becomes especially desirable in view of the report of Cori, Schmidt, and Cori (22) that the appearance of glucose-1-phosphoric acid is a consequence of a reversible enzymatic equilibrium. Adenylic acid is supposed to act as a coenzyme in both esterification and hydrolysis. The polysaccharide synthesized in the reaction is assumed not to be identical with glycogen, and gives a blue color with iodine which, by the way, was observed and described by Bernard (12).

The findings of Macfarlane (87; *cf.* also 78), who comments on her own

[19] According to P. Liang (Ann. **521**, 216 (1936)), living (also "impoverished") yeast cells almost completely lose their power of fermentation when exposed to freezing at very low temperatures.

experiments as follows: "The coincident changes in total, labile, and organic (acid-soluble) phosphate could not be consistently related to the fermentation process," are without significance for the understanding of the initiating phase in living-cell fermentation.

Neuberg (106) was able to demonstrate a phosphorylation in living yeasts, with the exception that in his experiments it was not a carbohydrate but a mono-ester which served as substrate.

The fermentation of disaccharides has also a particular interest in this connection, since no conclusion has yet been reached concerning the course of degradation and the sequence of the phases in the fermentation of these important carbohydrates by yeast. Since only a few observations concerning the rôle of phosphorylation have been made, it is still undecided whether phosphorylation has to be included in the sequence, and, if so, at what stage.

As an assumption for the fermentability of disaccharides, there was formerly required, in general, the presence of specific carbohydrases in the enzyme system of the corresponding organism. Even recently, Armstrong and Armstrong (4), for example, express themselves as follows on this question: "maltose is fermented only by those yeasts which contain maltase and then not until hydrolysis has been brought about by the enzyme," and later "only those yeasts which contain lactase are capable of fermenting milk-sugar"; then they say with respect to cane sugar, "cane sugar is fermented by yeasts only after previous inversion with the invertase of yeast."

Through the researches of Willstätter with Oppenheimer (179) on lactose, and with Steibelt (160) on maltose the question of the possibility of the direct fermentation of bioses was investigated. They assumed the existence of particular zymases for maltose and lactose in yeast which ferment the above-mentioned composite sugars without requiring a preliminary split to monoses. Sobotka and Holzman (156) believed that they were able to confirm these experiments by the use of a particular American yeast which respires as well as ferments and grows at a pH of 2.5, and supposed, in consequence, the existence of a particular maltozymase. They assumed, like Hvistendahl (62), that a disaccharide phosphate (similar to the isolated hexose esters) existed as an intermediate phase in the direct course of biose fermentation. This assumption was, however, disproved experimentally by Baba (9). Wright (183), by investigating the degradation of lactose by *Streptococcus thermophilus*, also recently arrived at the conclusion that disaccharides can undergo direct fermentation. Leibowitz and Hestrin (77), in experiments on the fermentation of α-methylglucoside, report the same results *cf.* 151a. Living *Fusarium lini* Bolley is capable of direct fermentation of maltose without phosphorylation (123).

There remains only the discussion of the question of the fermentability of cellular or of added trehalose by living yeast, the investigation of which was carried out by many workers. One of these investigators, Myrbäck (102), had to interpret his results in a contradictory way. In 1936 he believed that he had proved a measurable carbon dioxide evolution, lying beyond the limits of experimental error, with top as well as bottom yeast. A repetition of the experimental work compelled him in the years 1937–1939 to take the opposite stand.

Investigations conducted in the author's laboratory (127a) indicate that added trehalose, in the course of its fermentation by living *Fusarium lini* Bolley, is not necessarily esterified by added inorganic phosphate nor is its dissimilation dependent upon preliminary hydrolysis, inasmuch as its rate of fermentation exceeds that of glucose.

VII. FUNDAMENTAL CONSIDERATIONS

It was the everlasting contribution of Buchner and Hahn, by their discovery of cell-free fermentation, that they not only settled the dispute between Liebig and Pasteur, whose views were so happily supplemented by the unitary theory of enzyme action of Traube,[20] but that they also left to posterity a means and a method which have proved indispensable, although not decisive, for the investigation of the sequence of carbohydrate degradation and the enzymology of the corresponding systems. In the course of later investigations a mass of observations and results have been obtained which were fundamentally determined in the case of cell-free fermentation from two viewpoints, and in the case of living-cell fermentation from a third viewpoint. In the first case, a disturbance of the ratio of the various components and an injury to the total enzyme system is automatically caused, whereby an accumulation and even a stabilization of the intermediate products can arise (Iwanow (1905); Harden and Young

[20] The crystallized systems of Sumner, Northrop, Anson, Dounce, and others can at present be regarded as an experimental confirmation of this theory. However, it seems questionable to wish to support the homogeneity of these crystals by the measurements of Svedberg (K. G. Stern: Enzymologia **5**, 191 (1938)). Apart from the contrary findings which were obtained by means of cryolysis (*cf.*, for example, O. M. von Ranke-Abonyi and F. F. Nord: Kolloid-Z. **58**, 198 (1932)), it must be unequivocally concluded from the discussions of Bawden and Pirie or of K. M. Smith (Nature **142**, 842–3 (1938)), that purified tobacco mosaic virus undergoes an aggregation, i.e., a change of particle radius when centrifuged. Independently, R. W. G. Wyckoff (Cold Spring Harbor Symposia Quant. Biol. **6**, 365 (1938)) expressed himself on the question of "molecular weight" as follows: "I do not think that we should talk much about the molecular weights of the macromolecules until all the necessary factors have been determined." *Cf.* R. W. G. Wyckoff: Ergeb. Enzymforsch. **8**, 4 (1939).

(1906)). On the other hand, by selective poisoning of parts of the enzyme system a tremendous quantity of factual material has been furnished which will certainly not easily lose its statistical value. In living-cell fermentation there occurred, because of the introduction of a reagent not akin to the whole system, the removal of a supposed intermediary product whereby the same was excluded from the reaction sequence. This method, arising from the demands of war, originated with Connstein and Lüdecke (18)[21] and consisted in the application of an earlier observation of Dumas (29), who had established in 1874 the possibility of influencing alcoholic sugar degradation by the presence of alkali sulfites.

All these methods can, under favorable conditions, lead to the isolation of compounds, or their derivatives, which in the normal course of events would not appear. It is, however, to be emphasized that the removal of an intermediary product, as well as the selective poisoning or injury to the enzymatic system, leaves open the possibility of disturbing other phases along with the actual change in the normal sequence.

Important objections to the unlimited application of the reaction sequence found in cell-free fermentations and other preparations (due to their chaotic state, in contrast to the living structures) to the realm of action of enzyme systems in undamaged cells are to be found in the field of biophysical chemistry. First, the degree of dispersion (121, 125; *cf.* also 122) of the carrier of dualistic enzyme systems (*loc. cit.*) can change under the influence of various factors. Over and above this is the question whether portions of a sequence of phases in a cell-free or artificial enzyme system constitute a mirrored image of the actions of an enzyme system of living cells. In considering an answer to this question, account must be taken of the important findings of molecular anatomy, which show that the disperse particles of the various protoplasmic substances of cells have widely differing pH values. In other words, it is possible in the same cell for acid as well as alkaline particles to exist side by side *without* immediately neutralizing or precipitating each other (158; *cf.* also 167). Therefore, according to Spek (159),

"bei einem in gesetzmässiger Weise lokal zwischen 5.0 and 8.0 (oder darüber hinaus) variierenden pH schon viel eher auch *sie* (die Enzyme) *an bestimmten Orten in der Zelle zu höchster Aktivität gesteigert, an andern völlig wirkungslos werden können.* Die Frage der Lokalisation der Fermente und die Beschaffenheit ihrer kollidalen Träger wird jetzt von neuer Seite her sehr aktuell."[22] (See also page 442.)

---

[21] On the basis of the reports of Dr. Alonzo Taylor, the American Under-Secretary of Agriculture, who was in Germany in the summer of 1916, this procedure had already been tested in in the United States in the beginning of 1917, in order to determine whether it was suitable for technical use.

[22] Italics inserted by author.

From the latest advances in our knowledge of the surfaces within the cells (36),[23] we know that many of them possess particular properties and that the correct understanding of reactions stands in the closest connection with the properties of colloidal structure. Here physical and organic chemistry meet. No doubt, by an experiment carried out *in vitro* this intracellular arrangement of things is disturbed.

Finally, according to the measurements of Potter (137; *cf.* also 138), during the fermentation of sugar by the living cell 8 coulombs of electrical energy are liberated. This liberation is independent of the pH of the medium as long as the efficiency of the organism is not injured.

There arises now the occasion to indicate the possibility that the reactions of a cell-free fermentation constitute only fragments of the total reaction mechanism. Besides these, the details of the energy conversions, such as growth, energy transport,[24] and energy transformation, are of at least the same significance for the reactions of the living organisms as the stoichiometrically perceptible processes. Nevertheless, the clarification of the sequence of phases which is effected by the enzyme systems separated from the cells is a meritorious work. However, the attempt to interpret the action of living-cell systems only with the aid of the clarification mentioned before has a justification which borders on speculation.[25]

A further important experimental basis for this conception lies in the mode of action of the amino acid oxidase discovered by Krebs (73), the prosthetic group of which was isolated by Warburg and Christian (170), and the carrier of which was isolated by Negelein and Brömel (104). The amino acid oxidase occurring in tissue slices of kidney and liver is capable of oxidatively deaminating *l*- or *d*-amino acids. If, however, the tissue is injured by grinding, by drying, by octyl alcohol, or by hydrocyanic acid, then only the "unnatural" amino acids are attacked while the natural amino acids are no longer deaminated.

The living cell is thought of in this relation as an "organization." If the interpretation and presentation of a living system with the help of a chemical and physical terminology be regarded as the ultimate goal of biochemistry, then it must be clear that thereby the attempt is made to

[23] With regard to the properties of the outer surfaces of the cell, *cf.* E. N. Harvey and J. F. Danielli: Biol. Rev. **13**, 319 (1938).

[24] Even Meyerhof (*Handbuch der Physik*, Vol. 11, p. 249 (1926)) once expressed himself on these questions as follows: "Schliesslich sei noch einer wichtigen Bedeutung der freiwilligen Stoffwechselreaktionen, insbesondere der Oxydation, gedacht, nämlich die *Energie* für unfreiwillige chemische Vorgänge, insbesondere 'Synthesen', zu liefern." Compare also L. Algera: Rec. trav. botan. néerland **29**, 37 (1932).

[25] About ten years ago Hill concluded his address dedicated to the memory of Mond with the following statement: "It is dangerous to speculate too far, but it is foolish not to speculate at all."

penetrate into a field wherein chemical reactions are bound in a complexity which was not envisioned in the development of classical chemistry. The chemical (and still less, the stoichiometric) manner of expression does not possess a means of giving an exhaustive expression of this more comprehensive realm. Nevertheless, by the indication of this difference the view should not be taken that the one is of itself more important than the other. We must here remind ourselves of an earlier statement of Kögl (71), "... dass das Zustandekommen eines normalen Gewebes letzten Endes grossere Rätsel birgt als jenes eines chaotischen Zellgefüges."

Therefore, another still unsolved problem remains,—to find the way which will enable us to draw from the overlapping of the effects of Buchner's preparations and of the living cells, completely valid conclusions concerning the total activity of the latter.

How sharp the dispute of opinion is in the evaluation of the probative strength of the experimental material at hand may be impressively illustrated by a repetition of the following evaluation, written by Meyerhof in 1937 (96) in defense of his hypothesis:

> "Ob es in bestimmten Zellarten anaerobe Wege des Kohlenhydratabbaues gibt, die nicht über die hier beschriebenen phosphorylierten Intermediärprodukte führen, bleibe dahingestellt. Beweise dafür liegen nicht vor."

A year later Fink and Krebs (38) expressed themselves as follows on the same question:

> "Diese Gleichung (namely, that of Gay-Lussac) gestattet uns, mit Sicherheit die höchstmöglichen *Ausbeuten* an Alkohol und Kohlensäure aus Zucker anzugeben, wenn auch der *Mechanismus* dieser chemischen Umwandlung heute noch seiner endgültigen Lösung, namentlich in bezug auf die vitale Gärung harrt."

How slightly justified are the categorical explanations of Meyerhof, can, furthermore, be seen from a communication of Deuticke and Hollmann (25), who, on the basis of their experiments, draw the conclusion that, in contrast to the proceedings in the structurally destroyed, enzymatically incomplete organ, the carbohydrate degradation in the *intact* muscle proceeds less through the Harden and Young ester than through hexose monophosphate.

Moreover Shorr and Barker (151b) stated recently: "Disruption of cell structure may result in the liberation of highly reactive carbohydrate systems no longer under control by the intact cellular organisation. This fact should be borne in mind when interpreting data derived from tissue mince and extending them by analogy to the living cell."

Another staunch exponent of the conformity of the course of reaction in

the case of yeast juices and living yeast cells is Macfarlane (88), who lately even admits,

"Yet there are a number of differences between the cell and the cell-free extract which may be summarized as follows: fresh yeast does not ferment added glycogen or hexosediphosphate; the rate of fermentation of sugar is not increased by inorganic P, cozymase, or arsenate; there is no stoichiometric relationship between the CO_2 evolved and the actual decrease in inorganic P."

It is furthermore to be considered that a number of native proteins have been crystallized. In contrast, however, no crystallized *denatured* proteins have thus far become known. If we attempt to bring this conclusion into agreement with the clarification, just recently begun (on the basis of the work of Mirsky and Anson (3; also 185 and 93), for example) of the processes in the case of various types of denaturation of proteins, brought about in several ways, then the fact is not to be overlooked that, at least in the case of the known crystallized enzymes, regarded as unitary systems, the decrease in activity can be proven to run parallel to the increased denaturation. If this observation be transferred to the carrier proteins of dualistic enzyme systems, then it would be conceivable that in the enzyme systems (e.g., in maceration juices and others), in consequence of the possible dissolution or loosening of the various types of valence bindings of the carrier proteins during the preparation, a change in the specificity of the latter occurs which is capable of causing an extensive influence on the activities, or of eliminating parts of the total system. For these reasons, also, it is not conclusive if, from the enzymatic behavior of structurally destroyed systems, which can be more or less incomplete, forceful conclusions as to the qualitative actions of the parent systems within the living cells are drawn.[26]

VIII. MECHANISM OF FERMENTATION BY FUSARIA

The findings and their explanations selectively reported in section V are so contradictory that it seems impossible to reduce them at present in a constructive sense to a common denominator. In order, therefore, to be justified in drawing at least analogous conclusions, an attempt had to be made to bring forward for comparison a *living*-cell system or its metabolism, whose stoichiometrically conceivable effects are comparable to those of yeast. This is found in a group of polycellular fungi, the fusaria: *Fusarium lini* Bolley, *Fusarium oxysporum*, and *Fusarium graminearum*

[26] Compare also in this connection F. Hofmeister, *Die chemische Organisation der Zelle*, p. 8 (Braunschweig, 1901): ". . . so wenig der Biochemiker durch chemische Analyse einer zertrümmerten Taschenuhr deren regelmässigen Gang erklaren könne, ebensowenig sei von der chemischen Untersuchung des toten und zertrümmerten Protoplasmas eine Aufklärung über dessen Lebenserscheinungen zu erwarten."

Schwabe (*Gibberella saubinettii*), whose biology (182) has been very fruitfully studied. Concerning the biochemical effects of their enzyme systems only loose stoichiometric conclusions were known up to a few years ago.[27]

Fusaria cause a genuine alcoholic fermentation of hexoses as well as of pentoses. The ratio between carbon dioxide and ethyl alcohol in the case of the fermentation of hexoses is in almost complete agreement with that of a typical yeast fermentation. The ratio of carbon in the alcohol to the carbon in the carbon dioxide is 1 to 1 in the case of pentoses, in contrast to the ratio of 2 to 1 if the fungus grows on glucose. Accordingly, they possess zymases, phosphateses, and phosphatases, and an aeroglucosedehydrase and aeropentosedehydrase. They have a very powerful dehydrase system which enables them, for example, to utilize ethyl alcohol or polyvinyl alcohol as the only source of carbon and to dehydrogenate these with the formation of carbon dioxide.[28] They fer-

[27] A presentation by F. F. Nord and collaborators of the actual status of our knowledge of the mechanism of alcoholic fermentation and the other enzymatic effects of fusaria is to be found in Ergebnisse der Enzymforschung, Vol. 8 (1939), in the Biochemische Zeitschrift (1936–38), in the Berichte der deutschen chemischen Gesellschaft (1938), and in Chemiker-Zeitung (1938). Reports on the physiology of individual fusaria have been made by Y. Tochinai (J. Coll. Agr. Hokkaido Imp. Univ. **14**, 171 (1926); Ann. Phytopath. Soc. (Japan) **1**, part 3 (1920); Trans. Sapporo Nat. Hist. Soc. **8**, 1 (1920)), by H. H. Hochapfel (Zentr. Bakt. Parasitenk. II **64**, 174 (1925)), in the dissertation of G. Luz or H. Grossmann (Zürich, 1934), and by S. Medvedeva (Compt. rend. acad. sci. (U.R.S.S.) **15**, 503 (1937)).

[28] The complete phase sequence for the dehydrogenation of alcohols by means of *Fusarium lini* Bolley up to the stage of carbon dioxide can at present be formulated as follows:

(*a*) $C_2H_5OH \xrightarrow{+O_2} H_2O + CH_3COOH$

(*b*) $2CH_3COOH \xrightarrow{-H_2} HOOCCH_2CH_2COOH$

(*c*) $HOOCCH_2CH_2COOH \xrightarrow{-H_2} HOOCCH{=}CHCOOH$

(*d*) $HOOCCH{=}CHCOOH \xrightarrow{+H_2O} HOOCCH_2CHOHCOOH$

(*e*[1]) $HOOCCH_2CHOHCOOH \xrightarrow{-CO_2} CH_3CHOHCOOH$

(*e*[2]) $$2HOOCCH_2CHOHCOOH \rightarrow \underset{\displaystyle CH_3COCOOH}{\underset{\downarrow -CO_2}{HOOCCOCH_2COOH}} + \underset{\displaystyle HOOCCHOHCHOHCOOH}{\underset{\downarrow +H_2O}{HOOCC(OH){=}CHCOOH}}$$

In accordance with considerations presented in the literature (R. Sonderhoff: Ergeb. Enzymforsch. **3**, 163 (1934); K. Bernhauer: Ergeb. Enzymforsch. **3**, 216 (1934)) and undisputed experimental evidence, there first takes place a dehydrogenation of the succinic acid by way of fumaric acid to malic acid. The malic acid itself represents the branching point for the further course of reaction. It may lead (*a*) after de-

ment dextrins as well as disaccharides, of which, for example, maltose can be fermented without phosphorylation, directly and indirectly. They have a very powerful catalase system; their cell multiplication is promoted by hydrocyanic acid, and their dehydrases are not inhibited by this reagent. In comparison to *B. coli* or to yeast, the quantity of cellular phosphorus donors (adenosinetriphosphoric acid or muscle adenylic acid) present therein is small.

In contrast to yeast, these systems, however, are advantageous in that, since their cellular substrates are apparently not so easily capable of mobilization, the stoichiometrically conceivable part of the course of the reaction is slow and correspondingly more extended than in the case of the sequence in living yeasts. This afforded a possibility of obtaining an insight into the first phases of the dissimilation of the carbohydrates, which phases in the case of the undamaged yeast cells forced the adoption of the usual detours. In contrast to the conclusions of Meyerhof in the case of alcoholic fermentation by means of yeast juices (*cf.* page 442), it was shown with living fusaria, as well as with dried preparations, that the disappearance of inorganic[29] phosphorus during the fermentation of hexoses begins, on the average, 2 to 5 days after the onset of the fermentation and *after* the evolved carbon dioxide is obtained. At appropriate phosphate concentrations ribose and arabinose, as well as xylose, can be phosphorylated in the later course of the dissimilation. Added organic phosphorus donors have, in alcoholic fermentation with fusaria, only the function of cell regenerators. An accumulation of such donors could indeed be established in the micelle, but under normal conditions it was without essential influence on the kinetics of carbohydrate degradation. The hastening of the enzymatic reaction sequences, established under special conditions after addition of adenylic acid, was not necessarily caused by a phosphate transfer to the substrate, *since it was also observed*

carboxylation to lactic acid, and (*b*) after dehydrogenation as keto oxalacetic acid to pyruvic acid, and as an enol compound to tartaric acid.

Besides the identification and isolation of acetic, succinic, and malic acids in the course of the dehydrogenation of alcohols by means of the enzyme system of fusaria, lactic acid could also be qualitatively, and tartaric acid quantitatively, found as products of stabilization. There is, therefore, a restricted agreement with the findings of Chrzaszcz and Tiukow (Biochem. Z. **229**, 355 (1930)). But in order to avoid a disarrangement of the enzymatic course of the reaction and an unphysiological accumulation of transient products, we did not avail ourselves of the possibility of disturbing the reaction cycle (compare F. F. Nord: Naturwissenschaften **24**, 763 (1936)). Meanwhile a rapid and accurate analysis for the quantitative estimation of small amounts of succinic acid has been developed in this laboratory (45a).

[29] In these experiments special care was taken to avoid the presence of sodium fluoride as an impurity. *Cf.* A. Harden: Nature **134**, 101 (1934).

in a corresponding manner if alcohol was the substrate.[30] Accordingly, there are cell systems in which, during the course of alcoholic sugar decomposition, the degradation does not set in with phosphorylation. Correspondingly, there exists a fundamental agreement with the mechanism of carbohydrate metabolism in the case of brain tissue, of tumors (5, 44; *cf.*, to the contrary, 97), of chicken embryos (*loc. cit.*), and in the case of kojic acid formation by *Aspergillus tamarii* Kita (46) in which the phosphorus-free degradation is supreme. The phosphorylating degradation arises there also, as with fusaria (here also in the case of pentoses), only much later and to a quantitatively insignificant extent. The cellular physiological and evolutionary mechanical causes of this agreement might, however, be different for the systems mentioned. Whether the phosphorus-free course of degradation leads through a form of methylglyoxal can, at the moment, not be answered. Let it be noted here, however, that *Fusarium lini* Bolley is capable of using hydroxymethylglyoxal (62a, 126a), in the monomeric as well as in the trimeric form, as a source of carbon (118).

The present status of these researches gives further clues to the view that the application of experiments, correct in themselves, which were carried out with the help of cell-free enzyme systems, to the sequences of reaction of living cells is permissible only with the generally omitted limitation that they constitute, at best, only a part of the totality of reactions proceeding codependently in the living cell. Hence it is questionable whether all known "cell-free" reactions can demand a place in the reaction sequence in the living cell.[31]

However, using the wet crushing mill of Booth and Green (13a), an active cell-free juice was prepared from *Fusarium lini* Bolley which, thus far, seems to furnish an enzyme system which does not suffer the usual distortions and deficiencies of maceration juices obtained from yeast (180a).

In no case should the fact be overlooked that fusaria thus far constitute the only systems in which, in contrast to yeast, it has been possible in the latter course of the alcoholic fermentation with their *living* cells to establish a phosphorylation of the added hexoses or pentoses.

The discovery of the aeroglucosedehydrase and aeropentosedehydrase stands in good agreement with the supposition, already expressed by Boysen-Jensen (14), that various organisms can oxidize sugars. We have established the appearance of the corresponding acids. A splitting of the carbon chain, therefore, did not occur. It is interesting to see, and ought to be remembered, that the carbohydrate dissimilation can occur under

[30] This assumption is valid as long as the enzymatic formation of ethyl phosphates is not established.

[31] Compare also the discussion of the papers of Cori or Goddard in the Cold Spring Harbor Symposia on Quantitative Biology, Vol. 7 (1940).

the influence of fusaria in three ways: namely, (*a*) by oxidation, (*b*) by splitting of carbon chains, and (*c*) by the detour of phosphorylation.

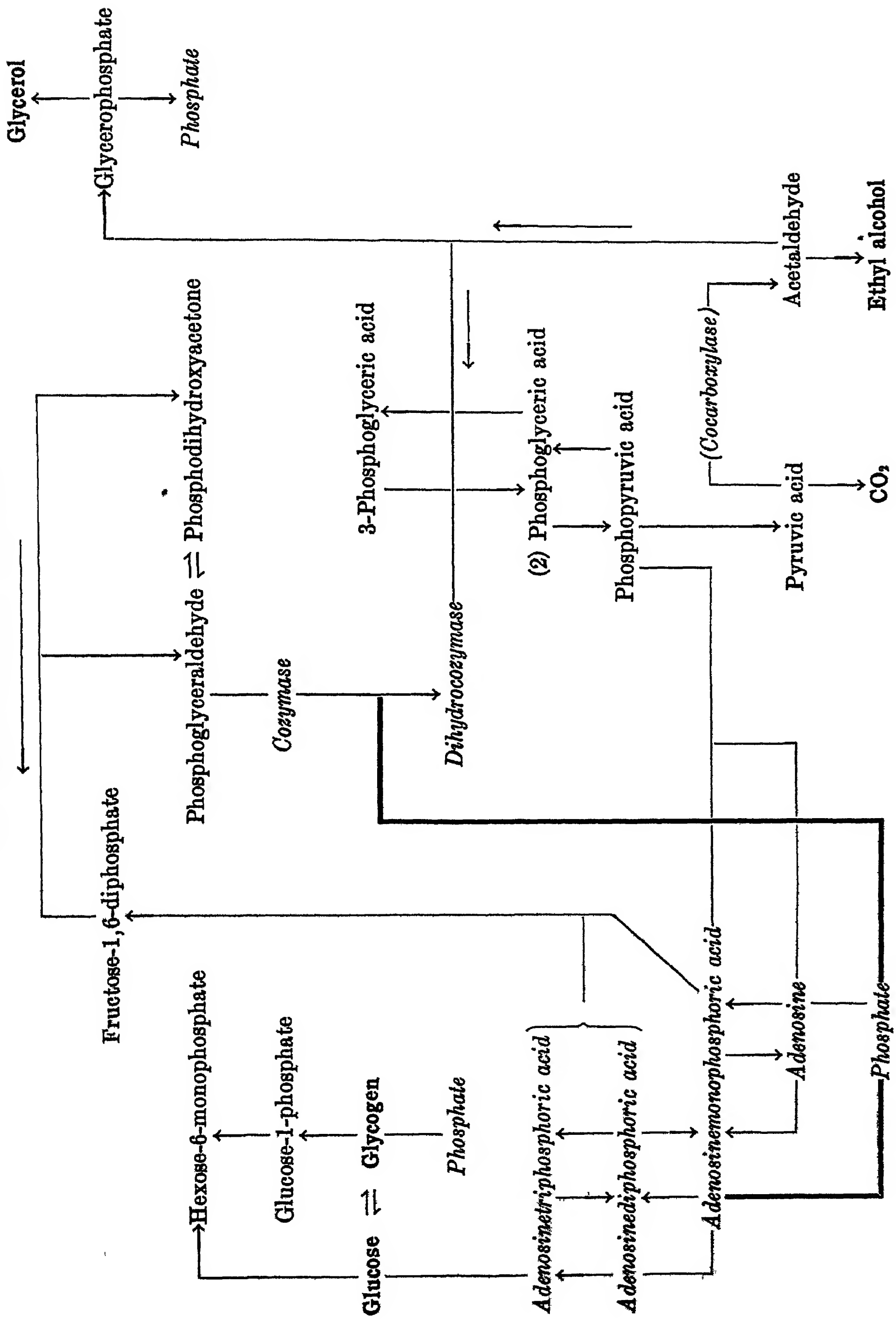
IX. PROCESSES IN THE COURSE OF CONVERSIONS WITHIN THE C_3 SERIES

If now the attempt is undertaken to present diagrammatically the pretended sequence of phases of alcoholic fermentation, starting from the previously mentioned stage of hexosemonophosphoric acid (*cf.* page 434), then we must keep in mind that here, as well as in the living yeast, 3-phosphoglyceric acid (114, 129; *cf.* also 9, 141, 168a) has a key position. Phosphoglyceric acid is formed after the previously mentioned hexosemonophosphoric acid has been converted into a di-ester, wherein adenylpyrophosphoric acid serves as a phosphate donor for the formation of the diphosphate. Phosphoglyceraldehyde or phosphodihydroxyacetone is formed in the presence of aldolase from the fructose-1,6-diphosphate. In the presence of cozymase (see page 460), phosphate, and adenosinediphosphoric acid, there arises from the phosphoglyceraldehyde the above-mentioned 3-phosphoglyceric acid ($\rightleftharpoons$ 2-phosphoglyceric acid), adenosinetriphosphoric acid, and dihydrocozymase. From the 2-phosphoglyceric acid there further arises phosphopyruvic acid, which with adenosine gives adenylic acid and pyruvic acid.[32] The pyruvic acid is split by carboxylase into acetaldehyde and carbon dioxide, and the acetaldehyde is stabilized into ethyl alcohol, whereby the cozymase is regenerated from the dihydrocozymase. According to a proposal by Parnas (132), these steps can be summarized in the sequential picture of alcoholic fermentation given in table 1.

[32] The conversion of phosphoglyceric acid to pyruvic acid can, according to Neuberg and Kobel (Biochem. Z. **272**, 459 (1934)) also be accomplished in a purely chemical way by removal of water, using potassium pyrosulfate. The fermentability of pyruvic acid found by O. Neubauer (*loc. cit.*) was originally denied by C. Neuberg and A. Hildesheimer (Biochem. Z. **31**, 173 (1911)). They wrote at the time: "Aus diesen Versuchen folgt, dass die freie Brenztraubensäure *nicht*, wohl aber ihre löslichen Alkali- und Erdalkalisalze mit Hefe 'gären'". This observation might, according to present knowledge, be applied to added acid (present in the keto-form). From measurements of adsorption spectra by Henri and Fromageot (Bull. soc. chim. [4] **37**, 852 (1925)), it is known that pyruvic acid *in vivo* occurs only in the easily fermentable enol-form. *Cf.* polarigraphic measurements of Müller and Baumberger (J. Am. Chem. Soc. **61**, 594 (1939)) and Zambotti and Ferrante (Boll. soc. ital. biol. sper. **14**, 372 (1939)). This observation could also be supported by the isolation of the so-called glucic acid, CHOH=CHCOOH, and also by its preparation (Nelson and Browne: J. Am. Chem. Soc. **51**, 830 (1929)). The latter arises from the action of weak calcium hydroxide on dilute solutions of glucose at 70°C. with the exclusion of air and greedily absorbs atmospheric oxygen with strong evolution of heat. According to Browne (Science **77**, 223 (1933)), the strong heat effect observed in hay fermentation can also possibly be traced to the oxidation of similar intermediate compounds.

TABLE 1

Phase diagram of alcoholic fermentation



Further details of the decomposition into the phosphorus-containing C_3 bodies and of the further decomposition of these C_3 compounds needs no special treatment. However, before the discussion of the biochemically and enzymologically most interesting parts of the reactions or of the sequences—namely, the mechanism of the hydrogenation and dehydrogenation with the help of cozymases, and the description of the action of carboxylases or cocarboxylases—is begun, the disproportionation of the C_3 bodies, already indicated on page 427, within the frame of the classical equation of Gay-Lussac,

$$C_6H_{12}O_6 = 2CO_2 + 2C_2H_5OH$$

should be elucidated. Müller-Thurgau and Osterwalder (100) observed that sulfurous acid added to fermenting sugar solutions reacts immediately with the acetaldehyde (133a) present in the mash. It is clear that this compound is to be regarded as the acetaldehyde sulfurous acid of Ripper (146),[33] the sodium salt of which has been known since the time of Bunte (17). In the course of comprehensive investigations on the hydrolysis of bound sulfurous acid, Kerp (67; *cf.* also 72) later established that the dissociation constant of acetaldehyde sodium sulfite (2.84×10^{-6}) bears a ratio to that of the corresponding glucose compound (311×10^{-3}) of approximately 1:90,000. On this basis were built up the experiments of Connstein and Lüdecke (18), which later found their biochemical interpretation in the light of the extensive investigations of Neuberg, Reinfurth, and Hirsch (112; *cf.* also 43). The above-mentioned acetaldehyde bisulfite compound, $CH_3CHOHOSO_2Na$, which can easily be split by hot soda solution with the reformation of acetaldehyde, accumulates in the mashes and can be separated. The corresponding hexose compound is, in the presence of water, practically completely dissociated. The relations can be reviewed by the following equation:

$$C_6H_{12}O_6 + Na_2SO_3 + H_2O = CH_3CHOHOSO_2Na + NaHCO_3 + C_3H_8O_3$$

The process in a salt-free solution is to be expressed by the following equation:

$$C_6H_{12}O_6 = CH_3CHO + CO_2 + C_3H_5(OH)_3$$

The ratio between glycerol and acetaldehyde is constant at any given instant.

If the fermentation be carried out in the presence of simple alkali salts which do not unite with acetaldehyde, only traces of the latter are found. In this case, the Cannizzaro dismutation sets in between two molecules of

[33] Compare, for the estimation of acetaldehyde, E. Parkinson and E. C. Wagner: Ind. Eng. Chem., Anal. Ed. 6, 433 (1934).

acetaldehyde and not between a molecule of acetaldehyde and a molecule of triosephosphoric acid. The acetaldehyde is found indeed in molecular proportions but as acetic acid or as ethyl alcohol.

$$2C_6H_{12}O_6 + H_2O = C_2H_5OH + CH_3COOH + 2CO_2 + 2C_3H_5(OH)_3$$

Two moles of glycerol correspond to each mole of acetic acid. That the last-named process, the "mixed" Cannizzaro reaction, if occurring between two different aldehydes, has, nevertheless, a fundamental significance could be shown by model experiments (116; 24) and *in vitro* (171). On the above basis, Neuberg and von Grab (47) were able to sustain the results of Fernbach and Schoen (37) by realizing, through the addition of β-naphthylamine to a fermenting cell-free sugar solution, the Döbner synthesis (28) of α-methyl-β-naphthocinchoninic acid and therewith isolating the transient pyruvic acid.[34]

$$CH_3COCOOH + CH_3CHO + C_{10}H_7NH_2 = C_{15}H_{11}NO_2 + 2H_2O + H_2$$

Using certain crystalline proteins in conjunction with dihydropyridine nucleotide, a triose or triosephosphate, it was demonstrated by Negelein and Broemel (104) that it is the protein of the reducing "Gärungsferment" which catalyzes the formation of glycerol.

By diminishing the enzyme concentration and simultaneously the concentration of the coenzymes by the introduction of a substance causing plasmolysis (toluene) (111), it was possible with top and bottom yeasts to change the alcoholic sugar dissimilation into a lactic acid fermentation (8). This was carried out with the aid of glutathione at a certain stage of the fermentation.

These observations, while in agreement with the findings of Aubel and

[34] A review of fifty different methods of determining pyruvic acid is given in a work of Wendel (J. Biol. Chem. **94**, 717 (1932)). A method which helps in determining 2γ in 10 cc. (corresponding to a dilution of 1:5,000,000) of blood pyruvate with an experimental error of 1.5 per cent was reported by G. D. Lu (Biochem. J. **33**, 249 (1939)). *Cf.* also Fromageot and Desnuelle: Biochem. Z. **279**, 174 (1935)). Here also attention may be drawn to the results of the investigation by C. V. Smyth (J. Biol. Chem. **125**, 635 (1938)) of the utilization of pyruvic acid by bakers' yeast. A remarkably high respiratory quotient was observed. The oxygen consumption and the loss of carbon dioxide were not equimolecular and did not correspond to the equation laid down for the decomposition. On the other hand, the acid caused no increase in the fermentation of reserve carbohydrate. All signs pointed to the appearance of strongly reducing substances, such as fats or fatty acids, which are indispensable as sources of energy. Without giving results on the fats or fatty acids formed, it could be concluded from the petroleum ether extract that the yeast under aerobic conditions can take up two to four times as much pyruvic acid as under similar anaerobic conditions. Concerning the composition of yeast fats compare G. Weiss: Biochem. Z. **243**, 269 (1931).

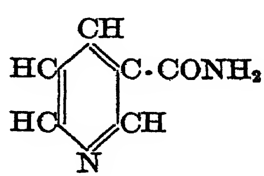
Simon (6), do not, so far as the question whether methylglyoxal should be considered as a primary transformation product in a fermentation by living cells, possess conclusive force since, according to Lohmann (85), the ketone-aldehyde mutase requires glutathione as a coenzyme. Enzyme systems containing the zymase complex which, however, are free from glutathione are not capable of attacking added methylglyoxal.

It is difficult not to see in these equations and findings a realization of the far-seeing statement of Pasteur (133):

> "Aujourd'hui, il faut comprendre, au contraire, que l'équation d'une fermentation est essentiellement variable avec les conditions dans lesquelles elle s'accomplit, et que la recherche de cette équation est un problème aussi compliqué que celui de la nutrition chez un être vivant."

X. COZYMASE AND ITS MODE OF ACTION

As has already been shown, the first oxido-reduction processes in the sequence given in table 1 set in after the splitting of the hexose diphosphate. Here is visible the influence of a dehydrogenating enzyme which also participates in the formation of lactic acid from pyruvic acid as well as in the formation of alcohol from acetaldehyde. In its action the system has been known since the discovery of the erstwhile coenzyme, the present cozymase system, by Harden and Young (51). The Harden–Young system consists of three components: (*1*) the phosphorylating coenzyme (cophosphorylase), which is an adenine nucleotide, and is either adenosinetriphosphoric acid (Lohmann) or diadenosinepentaphosphoric acid (Ostern); (*2*) cozymase (codehydrase I) which, according to the work of von Euler and Myrbäck (101) was regarded up to 1933 as an adenine mononucleotide. The molecular weight was given by von Euler, in agreement with the molecular weight of adenylic acid, as 350. The preparations gave on hydrolysis a content of 28 per cent adenine (168b). (*3*) The third constituent (Warburg and Christian) is a dinucleotide of molecular weight of approximately 800 (codehydrase II) and contains adenine as a purine base (approximately 17 per cent), the amide of nicotinic acid as a pyridine base (approximately 16 per cent),



which occurs therein in a quaternary linkage, and 13 per cent phosphorus. The elementary analysis (169) gives the empirical formula $C_{22}H_{32}O_{19}N_7P_3$,

which corresponds to 1 mole of adenine, 1 mole of pentose,[35] 1 mole of hexose, 3 moles of phosphoric acid, and 1 mole of nicotinic acid amide, minus 6 moles of water. Its active group is represented by the pyridine nucleus. By the reversible change pyridine $\rightleftharpoons$ dihydropyridine, this coenzyme transports hydrogen. On the basis of this discovery of the chemical mode of action of the last-named coenzyme, as well as the report of the analysis of nicotinic acid amide, the constitution of von Euler's cozymase could also be reinvestigated. The cozymase for this test was again obtained by von Euler from yeast and by Warburg from horse blood cells. In both cases we are dealing with hydrogen-transporting coenzymes,—"pyridine nucleotides,"—of which von Euler's compound, on the basis of the original determination of constitution by Warburg, is to be regarded as a *di*phosphopyridine nucleotide and Warburg's as a *tri*phosphopyridine nucleotide.[36] The close relationship of both coenzymes also made their mutual conversion possible, since codehydrase II on dephosphorylation can be converted into codehydrase I, and codehydrase I can be converted enzymatically and chemically into codehydrase II (*cf.* page 462). The redox potential of the diphospho system appears to lie in the vicinity of -0.26 volt (9a).

The fermentation test ("Gärtest") serves for the quantitative determination of the catalytic action of triphosphopyridine nucleotide. The substrate of the fermentation test consists of (*1*) hexosemonophosphoric acid, (*2*) phosphoric acid, and (*3*) acetaldehyde. In the test, the concomitant course of the following reactions can be weighed:

$$\text{1 hexosemonophosphoric acid} + \text{2 acetaldehyde} + \text{1 water} = \text{1 pyruvic acid} + \text{1 phosphoglyceric acid} + \text{2 alcohol} \qquad (1)$$

$$\text{1 hexosemonophosphoric acid} + \text{1 phosphoric acid} = \text{1 hexosediphosphoric acid} + \text{1 water} \qquad (2)$$

Here we are obviously concerned with a coupled reaction in which hexosemonophosphoric acid can be replaced by glucose or fructose. Hence no fermentation arises. If the hexosemonophosphoric acid be replaced by hexosediphosphoric acid, the fermentation becomes about ten times slower.

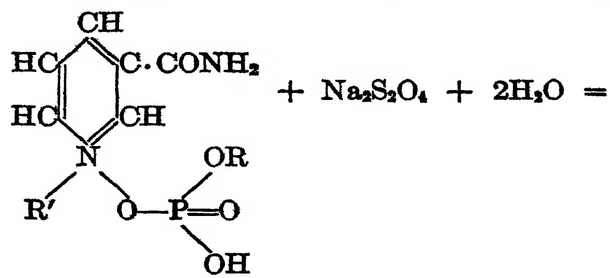
For the colorimetric determination (giving the order of magnitude) of nicotinic acid amide (in a cozymase preparation also), Karrer and Keller (65) employed a method which has as its basis the known reaction of

[35] The amount of pentose does not seem to be established beyond doubt. For the determination of small amounts of pentoses, *cf.* W. Mejbaum (Z. physiol. Chem. **258**, 117 (1939)).

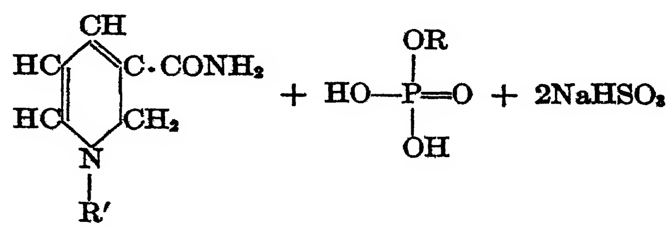
[36] Regarding nomenclature, see the proposal of F. G. Fischer (Ergeb. Enzymforsch. **8**, 187 (1939)).

pyridinium compounds with 2,4-dinitrochlorobenzene to give pyridinium salts. These are split by alkali to yellowish-red derivatives, $C_{11}H_9O_5N_3$, of glutaconaldehyde. The intensity of color of the solution obtained is measured by a "Stufen" photometer. A colorimetric method which serves to determine nicotinic acid and nicotinic acid amide in the presence of the codehydrases I and II with a limit of error of about 20 per cent was given by Euler and coworkers (34). Bandier and Hald (10) give an exact colorimetric determination of nicotinic acid which is based on the reaction of König (70). Cyanogen bromide at 70–80°C. is used, and the appearance of a color reaction with *p*-methylaminophenol is the basis of the determination. This method is to be preferred to both previously mentioned procedures. Dihydrocozymase is oxidized by the flavin enzyme, but not by oxygen or methylene blue.

Cozymase, which under physiological conditions behaves like a zwitterion anion (54), is reversibly reduced by hydrosulfite through the semiquinone stage (55), according to the following equation:



(R, R′ are sugar rests)



The mechanism shown proves that cozymase[37] acts as a hydrogen carrier somewhat in the sense of the equation

$$\text{cozymase} + \text{alcohol} = \text{dihydrocozymase} + \text{acetaldehyde}$$

The transport of the hydrogen follows only in union with a specific protein from yeast which has the rôle of mediating the oxido-reduction between

[37] So far as its preparation is concerned, compare P. Ohlmeyer: Biochem. Z. **297**, 66 (1938).

the hydrogenated codehydrases on the one hand and the yellow prosthetic groups on the other. In alcoholic fermentation, the loss of hydrogen follows through the donor system of the triosephosphoric acid which simultaneously becomes phosphoglyceric acid:

$$\text{glyceraldehydephosphoric acid} + \text{cozymase} \rightarrow \text{phosphoglyceric acid} + \text{dihydrocozymase}$$

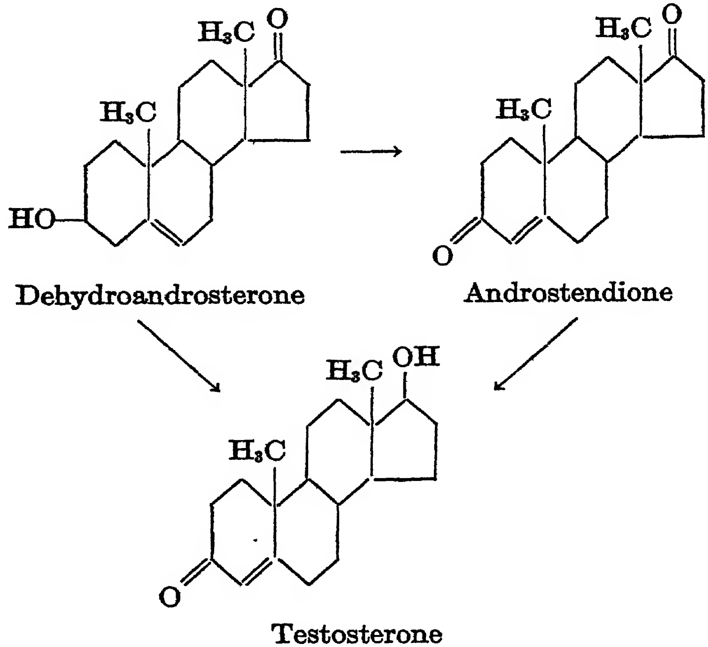
If, during the fermentation process a non-cellular hydrogen acceptor enters the arena in order to compete with acetaldehyde, there is present an example of the reaction known since the time of the discovery of the phytochemical hydrogenations by Lintner and von Liebig (80) and continued through the investigations of Neuberg,[38] Nord, Fischer, and Mamoli and Ercoli (90; *cf.* also 174).

An interesting contribution to the interpretation of the mechanism of these processes has been offered by the work of Fischer and Eysenbach (39). According to this work, the known hydrogen migration, which first reduces codehydrases, finally leads to the hydrogenation of ethylene linkages if the pH dependence of this reaction be considered. It is further important that the saturation of ethylenic linkages can also be carried out by enzyme solutions obtained from killed yeast. Here it is of interest that the addition of the unsaturated alcohol to the fermenting yeast juices carries with it a cessation of carbon dioxide evolution. A normally proceeding alcoholic fermentation does not, therefore, constitute a premise for the continuation of hydrogenation. If the fermentation be poisoned by fluoride ions, the speed of hydrogenation of an unsaturated alcohol is not changed by the removal of acetaldehyde as a concurrent acceptor. On the contrary, it was repeatedly observed that the reduction of cinnamaldehyde to cinnamic alcohol as well as the subsequent hydrogenation to saturated bodies was inhibited by the presence of fluoride. Fischer concludes therefrom that several reactions are capable of furnishing the hydrogen for the hydrogenation.

By union of biochemical hydrogenations and dehydrogenations, biochemical syntheses can be accomplished in a purely enzymatic manner. For instance, when dehydroandrosterone was subjected to the action of dehydrogenating bacteria and the reaction mixture was then added to fermenting yeast, the synthesis of the testicular hormone, testosterone, was accomplished. It is, therefore, conceivable that dehydroandrosterone

[38] The discussion of "carboligase" is omitted as, according to the findings of Dirscherl and Schöllig (Z. physiol. Chem. **252**, 71 (1938)), it does not exist. *Cf.* also B. Tanko and L. Munk: Z. physiol. Chem. **262**, 144 (1939).

constitutes an intermediate stage in the formation of testosterone in the organism.



XI. CARBOXYLASE AND COCARBOXYLASE

In 1910 Neubauer discovered the fermentability of pyruvic acid. The process corresponds to the equation

$$CH_3COCOOH \rightarrow CH_3CHO + CO_2$$

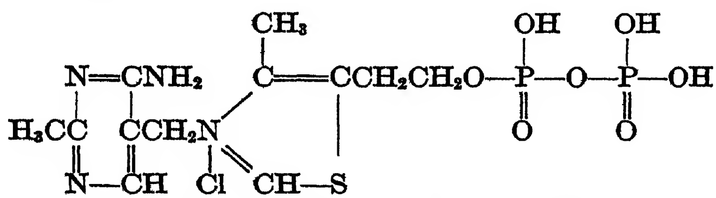
Although Neuberg (108) at first rejected[32] the finding of the fermentability of pyruvic acid, he later adopted the conception that pyruvic acid is an intermediary product of enzymatic carbohydrate degradation. According to his opinion, still regarded as correct, every zymase contains a carboxylase which splits carbon dioxide from pyruvic acid. If, according to Pasteur's hypothesis of *internal* respiration, the sugar is oxidized by oxidants which arise during the oxidation, e.g., acetaldehyde, then pyruvic acid arises from the sugar, while the acid in turn is hydrogenated to alcohol, corresponding to the equation:

$$1 \text{ hexose} + 2 \text{ aldehyde} = 2 \text{ pyruvic acid} + 2 \text{ alcohol}$$

From this equation, in the presence of carboxylase, pyruvic acid and acetaldehyde are eliminated.

The carboxylase mentioned here (holocarboxylase (2)) is composed of the actually active cocarboxylase (7) and the apocarboxylase of unknown constitution which exercises the function of a carrier. This discovery is related to the observation (84) that, by maintaining a definite hydrogen-ion concentration, it was possible to separate magnesium from zymase. Auhagen undertook the attempt to release further components from zymase in a weakly alkaline medium. Zymase is irreversibly damaged by phosphate solution at a pH of 8, while it can stand such a solution at pH of 7.8. An apozymase remains behind which can still be activated by boiled juice. Apozymase behaves likewise toward pyruvic acid after this treatment. Auhagen concluded therefrom that a component of carboxylase, for which he proposed the designation cocarboxylase, was separated from the apozymase.

Because of the accumulation of pyruvic acid in B_1 avitaminotic systems, there arose the hypothesis of a connection between cocarboxylase and vitamin B_1 (134). In fact, Lohmann and Schuster (86) were able to establish that cocarboxylase[39] is the diphosphate of vitamin B_1 with this structural formula:



(Molecular formula = $C_{12}H_{19}O_7N_4SP_2Cl$)

Cocarboxylase could be synthesized chemically by the action of phosphorus oxychloride on synthetic vitamin B_1 as well as enzymatically from vitamin B_1 and inorganic phosphate by washed yeasts (161, 33, 164, 152). The same cocarboxylase on a different specific carrier (pyruvodehydrase) is, according to Lipmann (81; *cf.* 82), active, in lactic acid bacteria also. The splitting off of carbon dioxide seems to him to be only a consequence of the dehydrogenation of a hydrated form of pyruvic acid. He also succeeded in proving that carboxylase and pyruvodehydrase possess the same coenzyme, and that this coenzyme constitutes the active group of pyruvodehydrase.

A further connection between carboxylase and cocarboxylase is found in the realm of the structure of model compounds. Langenbeck (74)

[39] It is of interest to note that, contrary to the explanation of Lohmann and Schuster (Biochem. Z. **294**, 188 (1937)), Hennessy and Cerecedo (J. Am. Chem. Soc. **61**, 179 (1939) concluded from their experiments that cocarboxylase (phosphorylated thiamin) is not biologically more active than free thiamin.

obtained in the course of his work compounds which showed extensive carboxylase-like activity and contained the amino group as an active group.

XII. APPLICATION OF ISOTOPES IN THE BIOCHEMISTRY OF ALCOHOLIC FERMENTATION

(*a*) Extensive experiments on the fermentation of sugar in heavy water were performed by O. Reitz (144) in consequence of the preliminary measurements of the kinetics of the alcoholic fermentation of sucrose and *d*-glucose by E. Pacsu (130). It could thereby be shown that the inclusion of deuterium into the methyl group takes place to the extent of about 33 per cent, which means that only one deuterium atom is included therein. The main quantity of the hydrogen included in the alcohol lies in the CH_2OH group. Fermentation alcohol formed in pure deuterium oxide would, accordingly, have the formula CH_2DCD_2OD. Independently of the conceptions that can be formed on the transitions involved herein, the decarboxylation of pyruvic acid leads to the conclusion that all the hydrogen atoms on the hydroxylated carbon atom of the alcohol become heavy.

The dependence of the velocity of fermentation on concentration shows that deuterium oxide reacts about half as fast as water in alcoholic fermentation.

(*b*) In the investigation of the uptake and loss of active phosphorus from nutrient solutions, the conclusion was reached with yeast that evidently no exchange of individual phosphorus atoms can occur between the cell and the medium. In the yeast, accordingly, phosphorus must be bound in an unexchangeable form, e.g., as adenylphosphoric acid (57).

REFERENCES

(1) Adler, E.: Arkiv Kemi, Mineral. Geol. **B12**, No. 38 (1937); **B12**, No. 54 (1938).
(2) Albers, H.: Naturwissenschaften **24**, 794 (1936).
(3) Anson, M. L.: In C. L. A. Schmidt's *Chemistry of the Amino Acids and Proteins*, p. 407. Charles C. Thomas, Springfield, Illinois (1938).
(4) Armstrong, E. F., and Armstrong, K. F.: *The Carbohydrates*, 5th edition, pp. 164, 170, 179. Longmans, Green and Company, London (1934).
(5) Ashford, C. A., and Holmes, E. G.: Biochem. J. **23**, 748 (1929).
(6) Aubel, E., and Simon, E.: Compt. rend. soc. biol. **114**, 905 (1933).
(7) Auhagen, E.: Z. physiol. Chem. **204**, 149 (1932); **209**, 20 (1932); Biochem Z. **258**, 330 (1933).
(8) Auhagen, E., Neuberg, C., and Auhagen, T.: Biochem. Z. **264**, 452 (1933); **268**, 247 (1934).
(9) Baba, T.: Biochem. Z. **273**, 207 (1934); **267**, 452 (1933).
(9a) Ball, E. G., and Ramsdell, P. A.: J. Biol. Chem. **131**, 767 (1939).
(10) Bandier, E., and Hald, J.: Biochem. J. **33**, 264 (1939); **33**, 1130 (1939).

(11) Benedict, S. R.: J. Biol. Chem. **76**, 457 (1928).
(12) Bernard, C.: *Leçons sur le diabète et la glycogenèse animale.* Librairie J.-B. Baillière et fils, Paris (1877).
(13) Bertrand, G.: Compt. rend. **124**, 1032 (1897).
(13a) Booth, V. H., and Green, D. E.: Biochem. J. **32**, 855 (1938).
(14) Boysen-Jensen, P.: Kgl. Danske Videnskab. Selskab Biol. Medd. **4**, 1 (1923).
(15) Buchner, E.: Ber. **30**, 117 (1897); **32**, 127 (1899).
(16) Buchner, E., and Antoni, W.: Z. physiol. Chem. **46**, 141 (1905).
(17) Bunte, H.: Ann. **170**, 309 (1873).
(18) Connstein, W., and Lüdecke, K.: German patents 298593 (1915); 298594-6, 12o (1916); 343321 (1917); 347604 (1917); Ber. **52A**, 99 (1919).
(19) Cori, C. F., and Cori, G. T.: Proc. Soc. Exptl. Biol. Med. **36**, 119 (1937).
(20) Cori, C. F., Colowick, S. P., and Cori, G. T.: J. Biol. Chem. **121**, 465 (1937).
(20a) Cori, G. T., Colowick, S. P., and Cori, C. F.: J. Biol. Chem. **124**, 543 (1938); **127**, 771 (1939).
(21) Cori, G. T., Colowick, S. P., and Cori, C. F.: J. Biol. Chem. **123**, 375 (1938).
(22) Cori, C. F., Schmidt, G. and Cori, G. T.: Science **89**, 464 (1939).
(23) Corran, H. S., Green, D. E., and Straub, F. B.: Biochem. J. **33**, 798 (1939).
(24) Davidson, D., and Bogert, M. T.: J. Am. Chem. Soc. **57**, 905 (1935).
(25) Deuticke, H. J., and Hollmann, S.: Z. physiol. Chem. **258**, 175 (1939).
(26) Dewan, J. G., and Green, D. E.: Biochem. J. **31**, 2327 (1937); **32**, 626 (1938).
(27) Dixon, H. H., and Atkins, W. R. G.: Sci. Proc. Roy. Dublin Soc. **14** [N. S.], No. 1 (1913).
(28) Döbner, O.: Ann. **242**, 265 (1887).
(29) Dumas, J. B.: Ann. chim. phys. [5] **3**, 57 (1874).
(30) Embden, G.: Klin. Wochschr. **12**, 213 (1933).
(31) Euler, H. von: In *Chemie der Hefe und der alkoholischen Gärung*, pp. 168–70. Akademische Verlagsgesellschaft, Leipzig (1915).
(32) Euler, H. von, and Myrbäck, K.: Z. physiol. Chem. **131**, 179 (1923) and subsequent papers.
(33) Euler, H. von, and Vestin, R.: Naturwissenschaften **25**, 416 (1937).
(34) Euler, H. von, and coworkers: Z. physiol. Chem. **256**, 225 (1938).
(36) Farr, Wanda K.: J. Phys. Chem. **42**, 1113 (1938).
(37) Fernbach, A., and Schoen, M.: Compt. rend. **157**, 1479 (1913); **158**, 1719 (1914).
(38) Fink, H., and Krebs, J.: Biochem. Z. **299**, 1 (1938).
(39) Fischer, F. G., and Eysenbach, H.: Ann. **529**, 87 (1937); **530**, 99 (1937).
(40) Fischer, H. O. L., and Baer, E.: Ber. **65**, 337, 1040 (1932).
(41) Gard, M.: Compt. rend. **194**, 1184 (1932).
(42) Gay-Lussac, L. J.: Ann. chim. phys. **76**, 245 (1810).
(43) Gehle, H.: Dissertation, Hannover, 1922.
(44) Geiger, A., and Magnes, T.: Biochem. J. **33**, 873 (1939).
(45) Goda, T.: Biochem. Z. **298**, 431 (1938).
(45a) Goepfert, G. J.: Biochem. J. **34** (1940), in press.
(46) Gould, B. S.: Biochem. J. **32**, 797 (1938).
(47) Grab, M. von: Biochem. Z. **123**, 69 (1921).
(48) Grüss, I.: Z. ges. Brauw. **27**, 686, 769 (1904).
(49) Haas, E.: Biochem. Z. **298**, 378 (1938).
(50) Harden, A., and Robison, R.: Proc. Chem. Soc. **30**, 16 (1914).
(51) Harden, A., and Young, W. J.: Proc. Chem. Soc. (London) **21**, 189 (1905).

(52) Harden, A., and Young, W. J.: Proc. Roy. Soc. (London) **B78**, 369 (1906).
(53) Harden, A., and Young, W. J.: Proc. Roy. Soc. (London) **B77**, 405 (1906).
(54) Hausser, I., and Kinder, E.: Z. physik. Chem. **B41**, 142 (1938).
(55) Hellström, H.: Z. physiol. Chem. **246**, 155 (1937).
(56) Hevesy, G. von: J. Chem. Soc. **1939**, 1219.
(57) Hevesy, G. von, Linderstrøm-Lang, K., and Nielsen, N.: Nature **140**, 725 (1937).
(58) Hofstetter, H., Leichter, H., and Nord, F. F.: Biochem. Z. **295**, 414 (1938).
(59) Holzapfel, L.: Kolloid-Z. **85**, 272 (1938).
(60) Holzapfel, L., and Nord, F. F.: Ber. **71**, 1217 (1938); Biodynamica **3**, No. 57 (1940).
(61) Hopkins, R. H., and Roberts, R. H.: Biochem. J. **29**, 919 (1935).
(62) Hvistendahl, B.: Angew. Chem. **46**, 335 (1933).
(62a) Hynd, A.: Biochem. J., **25**, 11 (1931).
(63) Iwanow, L.: Trav. soc. naturalistes St. Petersbourg **34** (1905).
(64) Jowett, M., and Quastel, J. H.: Biochem. J. **31**, 275 (1937).
(65) Karrer, P., and Keller, H.: Helv. Chim. Acta **21**, 463 (1938).
(66) Kerb, J.: Ber. **52**, 1798 (1919).
(67) Kerp, W.: Arb. Reichsgesundh. **21**, 220 (1904).
(68) Klem, A.: Hvalradets Skrifter **7**, 76 (1933).
(69) Kluyver, A. J., and Hoppenbrouwers, W. J.: Arch. Mikrobiol. **2**, 245 (1931).
(70) König, W.: J. prakt. Chem. **69**, 105 (1904).
(71) Kögl, F.: Z. Krebsforsch. **40**, 207 (1934).
(72) Kolthoff, I. M., and Furman, N. H.: Volumetric Analysis, Vol. I, pp. 177–88. John Wiley and Sons, Inc., New York (1928).
(73) Krebs, H. A.: Biochem. J. **29**, 1639 (1935).
(74) Langenbeck, W.: Chem.-Ztg. **60**, 953 (1936).
(75) Lebedew, A. von: Ann. Inst. Pasteur **26**, 31 (1912).
(76) Lehmann, I., and Martensson, E.: Skand. Arch. Physiol. **75**, 86 (1936).
(77) Leibowitz, J., and Hestrin, S.: Enzymologia **6**, 15 (1939).
(78) Lewitow, M. M.: Biochem. Z. **284**, 86 (1936).
(79) Lindner, P.: Z. Ver. deut. Zucker-Ind. **81**, 25 (1931).
(80) Lintner, C. J., and Liebig, H. F. von: Z. physiol. Chem. **72**, 449 (1911); **88**, 109 (1913).
(81) Lipmann, F.: Enzymologia **4**, 65 (1937).
(82) Lipmann, F.: Nature **144**, 381 (1939).
(83) Lipmann, F.: Festschrift für S. P. L. Sørensen, p. 317. Copenhagen (1938).
(84) Lohmann, K.: Biochem. Z. **237**, 445 (1931).
(85) Lohmann, K.: Biochem. Z. **254**, 332 (1932).
(86) Lohmann, K., and Schuster, P.: Naturwissenschaften **25**, 26 (1937).
(87) Macfarlane, M. G.: Biochem. J. **30**, 1369 (1936).
(88) Macfarlane, M. G.: Biochem. J. **33**, 574 (1939).
(89) Mahdihassan, S.: Biochem. Z. **226**, 203 (1930).
(90) Mamoli, L., and Ercoli, A.: Ber. **71**, 2698 (1938); Chimica e industria (Italy) **20**, 349 (1938); Ber. **72**, 190 (1939); Z. physiol. Chem. **261**, 287 (1939).
(91) Mathews, A. P., and Glenn, T. H.: J. Biol. Chem. **9**, 51 (1911).
(92) Medwedew, G.: Protoplasma **27**, 252 (1937).
(93) Meyer, K. H.: Proc. Roy. Soc. (London) **A170**, 68 (1939).
(94) Meyerhof, O.: Biochem. Z. **267**, 313 (1933).
(95) Meyerhof, O.: Biochem. Z. **273**, 84 (1934).

(96) MEYERHOF, O.: Ergeb. Physiol. biol. Chem. exptl. Pharm. **39**, 71 (1937).
(97) MEYERHOF, O.: Bull. soc. chim. biol. **20**, 1335 (1938).
(98) MEYERHOF, O.: Bull. soc. chim. biol. **20**, 1345 (1938).
(99) MEYERHOF, O., OHLMEYER, P., GENTNER, W., AND MEYER-LEIBNITZ, H.: Biochem. Z. **298**, 396 (1938).
(100) MÜLLER-THURGAU, H., AND OSTERWALDER, A.: Landw. Jahrb. Schweiz **1914**, 480.
(101) MYRBÄCK, K.: Ergeb. Enzymforsch. **2**, 139 (1933).
(102) MYRBÄCK, K.: Biochem. Z. **288**, 329 (1936); Svensk Kem. Tid. **49**, 24 (1937); **51**, 36 (1939).
(103) NEEDHAM, J., AND LEHMANN, H.: Biochem. J. **31**, 1210 (1937).
(104) NEGELEIN, E., AND BRÖMEL, H.: Biochem. Z. **300**, 225 (1939); **303**, 231 (1939).
(105) NEUBAUER, O.: Z. physiol. Chem. **70**, 350 (1910).
(106) NEUBERG, C.: Biochem. Z. **280**, 163 (1935).
(107) NEUBERG, C., AND EULER, H. VON: Biochem. Z. **240**, 245 (1931).
(108) NEUBERG, C., KARCZAG, L., AND HIRSCH, J.: Biochem. Z. **36**, 76 (1911); **98**, 141 (1919).
(109) NEUBERG, C., AND KERB, J.: Z. Gärungsphysiol. **1**, 114 (1912).
(110) NEUBERG, C., AND KOBEL, M.: Biochem. Z. **247**, 246 (1932).
(111) NEUBERG, C., AND NORD, F. F.: Biochem. Z. **67**, 12 (1914).
(112) NEUBERG, C., REINFURTH, E., AND HIRSCH, J.: Biochem. Z. **89**, 365 (1918); **92**, 234 (1918); **96**, 175 (1919); **98**, 141 (1919); **100**, 304 (1919); **105**, 307 (1920); **110**, 193 (1920); Ber. **52**, 1702 (1919); **53**, 1039 (1920).
(113) NILSSON, R.: Arkiv Kemi, Mineral. Geol. **A10**, No. 7 (1930).
(114) NILSSON, R.: Svensk Kem. Tid. **45**, 129 (1933).
(115) NILSSON, R., AND ALM, F.: Biochem. Z. **286**, 373 (1936).
(116) NORD, F. F.: Beitr. Physiol. **2**, 301 (1924); Chem. Rev. **3**, 65 (1926).
(117) NORD, F. F.: Nature **120**, 82 (1927); Ergeb. Enzymforsch. **2**, 23 (1933).
(118) NORD, F. F.: Protoplasma **2**, 303 (1927).
(119) NORD, F. F.: J. Phys. Chem. **31**, 867 (1927); *cf.* also Protoplasma **10**, 41 (1930).
(120) NORD, F. F.: Trans. Faraday Soc. **26**, 760 (1930).
(121) NORD, F. F.: Ergeb. Enzymforsch. **2**, 23 (1933).
(122) NORD, F. F.: Naturwissenschaften **24**, 481 (1936).
(123) NORD, F. F., AND ENGEL, W.: Biochem. Z. **296**, 153 (1938).
(124) NORD, F. F., AND FRANKE, K. W.: Protoplasma **4**, 547 (1928).
(125) NORD, F. F., AND LANGE, E. M.: Biochem. Z. **278**, 173 (1935).
(126) NORD, F. F., AND COWORKERS: Z. Elektrochem. **35**, 612 (1929); Protoplasma **10**, 41 (1930).
(126a) NORRISH, R. G. W., AND GRIFFITH, J. G. A.: J. Am. Chem. Soc. **60**, 1628 (1938).
(127) NOWINSKI, W. W.: Biochem. J. **33**, 978 (1939).
(127a) O'CONNOR, RITA C.: Biochem. J; **34** (1940), in press.
(128) OSTERN, P.: Enzymologia **3**, 5 (1937).
(129) OSTERN, P., AND GUTHKE, A. J.: Z. physiol. Chem. **248**, 155 (1937).
(130) PACSU, E.: J. Am. Chem. Soc. **56**, 245 (1934).
(131) PARNAS, I. K.: Ergeb. Enzymforsch. **6**, 57 (1937).
(132) PARNAS, I. K.: Enzymologia **5**, 172 (1938).
(133) PASTEUR, L.: *Études sur le bière*, p. 268. Gauthier-Villars, Paris (1876).
(133a) PELTIER, R.: L'aldehyde, sa formation et sa disparition dans la fermentation alcoölique, Soc. sci. phys. et nat. de Bordeaux, 1913.

(134) Peters, R. A.: Biochem. J. **30**, 2206 (1936).
(135) Pinkus, G.: Ber. **31**, 32 (1898).
(136) Porchet, B.: Ann. fermentations **4**, Part 7 (1938).
(137) Potter, M. C.: Zentr. Bakt. Parasitenk. II, **93**, 1 (1935).
(138) Potter, M. C.: *Bioelectric Potentials*, pp. 7, 38. University Press, Cambridge (1933).
(139) Przylecki, St. J. v.: Ergeb. Enzymforsch. **4**, 111 (1935).
(140) Ramsden, W.: Nature **142**, 1120 (1938).
(141) Rapoport, S.: Enzymologia **3**, 52 (1937).
(142) Raymond, A. L.: Proc. Natl. Acad. Sci. (U. S.) **11**, 622 (1925).
(143) Raymond, A. L., and Blanco, J. G.: J. Biol. Chem. **80**, 631 (1928).
(144) Reitz, O.: Z. physik. Chem. **175A**, 257 (1936).
(145) Remesow, I.: Bull. biol. méd. exptl. (U.S.S.R.) **3**, 606 (1937); Kausche, G., and Holzapfel, L.: Naturwissenschaften **28**, 62 (1940).
(146) Ripper, M.: J. prakt. Chem. [2] **46**, 428 (1892).
(147) Robison, R.: Biochem. J. **16**, 809 (1922).
(148) Sansum, W. D., and Woodyatt, R. I.: J. Biol. Chem. **24**, 346 (1916).
(149) Schäffner, A.: Z. physiol. Chem. **243**, 157 (1936).
(149a) Schäffner, A.: Naturwissenschaften **27**, 195 (1939).
(150) Schäffner, A., and Specht, H.: Z. physiol. Chem. **251**, 148 (1938).
(151) Schäffner, A., and Specht, H.: Naturwissenschaften **26**, 494 (1938).
(151a) Schultz, A. S., and Atkin, L.: J. Am. Chem. Soc. **61**, 295 (1939).
(151b) Shorr, E., and Barker, S. B.: Biochem. J. **33**, 1809 (1939).
(152) Silverman, M., and Werkman, C. H.: Enzymologia **5**, 385 (1939).
(153) Slator, A.: J. Chem. Soc. **89**, 133 (1906); **97**, 922 (1910).
(154) Smit, Jan: *Die Gärungssarcinen.* G. Fischer, Jena (1930).
(155) Smythe, C. V., and Gerischer, W.: Biochem. Z. **260**, 414 (1933).
(156) Sobotka, H., and Holzman, M.: Biochem. J. **28**, 734 (1934).
(157) Somogyi, M.: Proc. Soc. Exptl. Biol. Med. **24**, 320 (1927).
(158) Spek, J.: Ergeb. Enzymforsch. **6**, 1 (1937).
(159) Spek, J.: Kolloid-Z. **85**, 167 (1938).
(160) Steibelt, W.: Z. physiol. Chem. **115**, 211 (1921).
(161) Stern, K. G., and Hofer, J. W.: Enzymologia **3**, 82 (1937).
(162) Tait, A., and Fletcher, L.: J. Inst. Brewing **32**, 393 (1926).
(163) Tankó, B., and Robison, R.: Biochem. J. **29**, 961 (1935).
(164) Tauber, H.: J. Am. Chem. Soc. **60**, 730 (1938).
(165) Theorell, H.: Ergeb. Enzymforsch. **6**, 111 (1937).
(166) Tholin, Th.: Z. physiol. Chem. **115**, 245 (1921).
(167) Tschirch, A.: Helv. Chim. Acta **17**, 996 (1934).
(168) Turpin, R.: Compt. rend. **7**, 402 (1838).
(168a) Vercellone, A., and Neuberg, C.: Biochem. Z. **280**, 161 (1935).
(168b) Warburg, O.: Helv. Chim. Acta **19**, E82 (1936).
(169) Warburg, O., and Christian, W.: Biochem. Z. **287**, 291 (1936).
(170) Warburg, O., and Christian, W.: Biochem. Z. **298**, 150 (1938).
(171) Warburg, O., and Christian, W.: Biochem. Z. **301**, 221 (1939); **303**, 49 (1939)
(172) Werkman, C. H., Stone, R. W., and Wood, H. J.: Enzymologia **4**, 24 (1937).
(173) Wertheimer, E.: Fermentforschung **11**, 22 (1929); *cf.* also Z. angew. Chem. **51**, 739 (1938); Enzymologia **7**, 58 (1939).
(174) Wettstein, A.: Helv. Chim. Acta **22**, 250 (1939).
(175) Wieland, H., Claren, O. B. and Wille, F.: Ann. **492**, 183 (1932); **503**, 70 (1933).

(176) WIELAND, H., PROBST, O., AND CRAWFORD, M.: Ann. **536**, 54 (1938).
(177) WIELAND, H., AND WILLE, F.: Ann. **515**, 260 (1935).
(178) WIGGERT, W. P., AND WERKMAN, C. H.: Biochem. J. **33**, 1061 (1939).
(179) WILLSTÄTTER, R., AND OPPENHEIMER, G.: Z. physiol. Chem. **118**, 168 (1922).
(180) WILLSTÄTTER, R., AND ROHDEWALD, M.: Z. physiol. Chem. **247**, 272 (1937).
(180a) WIRTH, JOHN C., AND NORD, F. F.: Science **91–92** (1940), in press.
(181) WOHL, A.: *cf.* O. Lippmann in *Chemie der Zuckerarten*, p. 1891. Friedr. Vieweg and Sons, Braunschweig (1904).
(182) WOLLENWEBER, H., AND REINKING: *Die Fusarien.* Julius Springer, Berlin (1935).
(183) WRIGHT, H. D.: J. Path. Bact. **43**, 487 (1936); **45**, 117 (1937).
(184) WROBLEWSKI, A.: J. prakt. Chem. **64**, 1 (1901).
(185) WU, H.: Chinese J. Physiol. **5**, 321 (1931).
(186) YAMAMOTO, K.: J. Biochem. (Japan) **24**, 467 (1936).

No claim is made for the completeness of the above bibliography. For further references and historical background consult the following:

BARRON, E. S. G.: Physiol. Rev. **19**, 184 (1939).
BURK, DEAN: Occasional Publications of the American Association for the Advancement of Science **4**, 121–61 (1937).
FRANKENBURGER, W.: "Katalyse," in Ullmann's *Enzyklopädie der technischen Chemie*, 2nd edition, **6**, 436. Urban and Schwarzenberg, Berlin and Vienna (1930).
FRANKENBURGER, W.: *Katalytische Umsetzungen in homogenen und enzymatischen Systemen.* Akademische Verlagsgesellschaft, Leipzig (1937).
HARDEN, A.: *Alcoholic Fermentation*, 4th edition. Longmans, Green and Company, London (1932).
HILL, A. V.: Nature **123**, 723 (1929).
HOGNESS, T. R.: Proceedings of the Sixth Summer Conference on Spectroscopy, p. 31. New York (1939).
HOPKINS, F. G.: Lancet **1938**, 1147.
HVISTENDAHL, B.: *Das Zymophosphat und die alkoholische Gärung.* Aktiebolaget Tidning för Skaraborgs Län, Mariestad (1934).
KLUYVER, A. J.: *The Chemical Activities of Micro-Organisms.* University of London Press, Ltd., London (1931).
NEEDHAM, J.: *Order and Life.* Yale University Press, New Haven, Connecticut (1936).
NORD, F. F.: *Mechanism of Enzyme Action and Associated Cell Phenomena.* The Williams & Wilkins Company, Baltimore, Maryland (1929).
NORD, F. F.: *Zum Mechanismus der Enzymwirkung unter besonderer Berücksichtigung der Kryolyse.* F. Enke, Stuttgart (1933).
NORD, F. F.: Angew. Chem. **47**, 491 (1934).
NORD, F. F.: *Handbuch der Enzymologie.* Akademische Verlagsgesellschaft, Leipzig (1940).
ROBISON, R.: *The Significance of Phosphoric Esters in Metabolism.* New York University Press, New York (1932).
SCHOEN, M.: "Problèmes actuels de la chimie des fermentations," in Ann. fermentations **1**, 257 (1935).
SCHOEN, M.: *Faits nouveaux et nouvelles hypothèses dans la chimie des fermentations.* Masson & Cie., Paris (1939).

TAMIYA, H.: *Le bilan matériel et l'energêtique des synthèses biologiques.* Herman & Cie., Paris (1935).
TANKÓ, B.: *A biológiai cukurbomlások kémiája.* Debrecen (1937).
TRAUBE, M.: *Theorie der Fermentwirkungen.* Ferd. Dümmler's Verlag, Berlin (1858).
WETZEL, K.: Ergeb. Biol. **10**, 323 (1934).
WITZEMANN, EDGAR J.: Am. Naturalist **67**, 163, 1, 341 (1933); **68**, 150 (1934).
Cf. also the reviews in Annual Review of Biochemistry, Volumes 1–8 (1932–39), or Ergebnisse der Enzymforschung, Volumes 1–8 (Leipzig, 1932–39).

For methods see the following:
BERNHAUER, K.: *Gärungschemisches Praktikum,* 2nd edition. Julius Springer, Berlin (1939).

MICROFILM SETS OF PERIODICALS

The Committee on Scientific Aids to Learning, President James B. Conant of Harvard University, Chairman, has made a grant to cover the cost of making a microfilm master negative, on the most expensive film, of sets of volumes of scientific and learned journals.

This permits the non-profit Bibliofilm Service to supply microfilm copies at the sole positive copy cost, namely, one cent per page for odd volumes, or a special rate of half a cent per page for any properly copyable ten or more consecutive volumes.

The number of pages will be estimated on request to the American Documentation Institute, in care of the offices of Science Service, 2101 Constitution Avenue, Washington, D. C.

ERRATA

Volume 25, Number 2, October, 1939

Page 173, table 4:

For "K/K_u" substitute "10^5K".

For "8.1" (opposite *m*-Methoxybenzoic acid) substitute "9.0".

For "*p*-Hydroxycinnamic acid" and "*p*-Methoxycinnamic acid" substitute "*m*-Hydroxycinnamic acid" and "*m*-Methoxycinnamic acid".